

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2003, 17:13:34 : Search time 7709 Seconds
(without alignments)
11414.789 Million cell updates/sec

Title: US-09-601-168b-1
Perfect score: 2151
Sequence: 1 tgcgttgctgcgcctggc.....gtttgccagaaaaaaaaa 2151

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.ste.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.mu.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2151	100.0	2151	6	AX019507	AX019507 Sequence
2	2151	100.0	2151	6	AX057166	AX057166 Sequence
3	2151	100.0	2151	6	BD131778	BD131778 Human bet
4	2151	100.0	2151	9	AF129530	AF129530 Homo sapi
5	2151	100.0	2151	9	HSBTRCP	Y14153 Homo sapien
6	2030.4	94.4	2308	9	BC027994	BC027994 Homo sapi
7	1682.2	78.2	2175	10	AF081887	AF081887 Mus muscu
8	1663.2	77.3	1818	9	AF101784	AF101784 Homo sapi
9	1630.4	75.8	2081	10	BC003989	BC003989 Mus muscu
10	1567.2	72.9	1979	10	AF110396	AF110396 Mus muscu
11	1451.2	67.5	1712	10	AF099932	AF099932 Mus muscu
12	1442.8	67.1	1710	10	AF112979	AF112979 Mus muscu
13	1017.8	47.3	1671	5	XELSBTRCP	M98468 African cla
14	883.4	41.1	2134	9	AB033279	AB033279 Homo sapi
15	883.4	41.1	2252	9	AB033280	AB033280 Homo sapi
16	883.4	41.1	2274	9	AB033281	AB033281 Homo sapi
17	883.4	41.1	4230	9	AB014596	AB014596 Homo sapi
18	883.4	41.1	4395	9	BC026213	BC026213 Homo sapi
19	882.4	41.0	1653	9	AF176022	AF176022 Homo sapi
20	877.8	40.8	3874	10	AB093260	AB093260 Mus muscu
21	877.8	40.8	4030	10	BC034261	BC034261 Mus muscu
22	853.2	39.7	1970	10	AY038079	AY038079 Mus muscu
23	810.2	37.7	2530	5	BC045356	BC045356 Danio rer
24	757.8	35.2	67008	9	AP000252	AP000252 Homo sapi
25	757.8	35.2	100000	9	AP000212	AP000212 Homo sapi
26	757.8	35.2	149298	9	AP000031	AP000031 Homo sapi
27	757.8	35.2	340000	9	AP001711	AP001711 Homo sapi
28	742.6	34.5	2445	3	AB076893	AB076893 Clona int
29	742.6	34.5	2514	3	AK114399	AK114399 Clona int
30	676.4	31.4	2534	3	AV118898	AV118898 Drosophill
31	673.6	31.3	292639	2	AC112602	AC112602 Rattus no
32	673.2	31.3	2154	3	AF032878	AF032878 Drosophill
33	673.2	31.3	2367	3	AF222924	AF222924 Drosophill
34	673.2	31.3	3532	3	AF222923	AF222923 Drosophill
35	657	30.5	657	6	AX050273	AX050273 Sequence
36	518.8	24.1	1561	6	AX057148	AX057148 Sequence
37	512	23.8	934	5	XLUG3921	U63921 Xenopus lae
38	502.6	23.4	1827	5	XLUG3922	U63922 Xenopus lae
39	463.6	21.6	2983	10	BC008552	BC008552 Mus muscu
40	454	21.1	1443	3	AF339101	AF339101 Heteroder
41	453.2	21.1	2700	3	AF275253	AF275253 Caenorhab
42	386.2	18.0	1217	3	AK115763	AK115763 Clona int
43	333.2	15.5	108503	9	AL627424	AL627424 Human DNA
44	288	13.4	173634	3	AC009344	AC009344 Drosophill
45	288	13.4	235707	3	AE003733	AE003733 Drosophill

ALIGNMENTS

RESULT 1
AX019507
LOCUS AX019507 Sequence 1 from Patent WO938969. 2151 bp DNA linear PAT 07-SEP-2000
DEFINITION AX019507
ACCESSION AX019507
VERSION AX019507.1 GI:10043427
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Arenzana,S.F., Concordet,J.P., Kroll,M., Durand,H., Benarous,R. and Margottin,F.
TITLE Protein humane beta -trcp
JOURNAL Patent: WO 938969-A 1 05-AUG-1999;

ARENZANA SEISDEDOS FERNANDO (FR); CONCORDET JEAN PAUL (FR); INST
NAT SANTE RECH MED (FR); KROLL MATHIAS (FR); DURAND HERVE (FR);
BENAROUS RICHARD (FR); MARGOTTIN FLORENCE (FR); PASTEUR INSTITUT
(FR)

FEATURES

source

Location/Qualifiers

1..2151
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="ADNC codant pour la proteine humaine BtrCp"
70..1779
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07740.1"
/db_xref="GI:10043428"
/translation="MDPAEVLQEKALKFNNSSPERDNNQGEPPRKILPEKNSLRQTY
NSCARLCLNQSTVCLASTAMKTEVCYAKTKLANGSTSMIVPKORKLSAYEKEKELCV
KIFQWSESDQVEFVEHLISQMHYQGHINSYLKPLQDFITALPARGLDHIENI
LSYLDKSLCAELVCKEWRVTSQDMKWLIERVWRTSLWRGLAERGGQLFK
NKPPDGNAPNSFYRALYPKIIOIETIESNWRGSRHSLOIRHCRSETSKVYCLQYD
DQKIVSLGRDNTIKIMDKNTLECKRIILIGHTSVLCLOYDERVILITGSSDSVVRVNDV
NTGMDLNLHHCVALHLRFNNGMVTCKSDRSIAVMDWASPTDITLRLVLGHRAA
VNVDFDKYIVSAGSDTLKVWNTSCFVRLTNGHKGICLQYRDLVYSGSDN
TIRLWDIEGACLRLVGEHELVCIRPDNKRIVSGAYDQKIKVMDVLAALDPAPAG
TLCURTLVHSGRVRFLQDFEQIVSSSHDDTILIMDFLNDPAAQAEPSPRSRTVY
ISR"

BASE COUNT 628 a 467 c 513 g 543 t

CDS

BASE COUNT
ORIGIN

Query Match 100.0%; Score 2151; DB 6; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCGGTGGCTGGCGCTGGCCACCAAGGGGGCGCCGGGAGCGGACCGGACCGTGGCC 60
DB 1 TCGGTGGCTGGCGCTGGCCACCAAGGGGGCGCCGGGAGCGGACCGGACCGTGGCC 60
QY 61 TCGGGGATATGCGACCGCGGCGGCGGTGTGCAAGAGAGGACCTCAAGTTATGAAT 120
DB 61 TCGGGGATATGCGACCGCGGCGGCGGTGTGCAAGAGAGGACCTCAAGTTATGAAT 120
QY 121 TCCTCAGAGAGAGAGACTGTAATAATGGCGAACCCCTAGGAAGATAATACAGAGAAG 180
DB 121 TCCTCAGAGAGAGAGACTGTAATAATGGCGAACCCCTAGGAAGATAATACAGAGAAG 180
QY 181 AATTCATTAGACAGACATACAAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
DB 181 AATTCATTAGACAGACATACAAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
QY 241 TGTTTAGCAAGCACTGCTATCAAGACTGAGAAATGTGTGGCCAAACCAAACTTGCCAAT 300
DB 241 TGTTTAGCAAGCACTGCTATCAAGACTGAGAAATGTGTGGCCAAACCAAACTTGCCAAT 300
QY 301 GGCATTCAGATGATGTGCCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAA 360
DB 301 GGCATTCAGATGATGTGCCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAA 360
QY 361 AAGGAACGTGTGTCAATATCTTTGACAGGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTG 420
DB 361 AAGGAACGTGTGTCAATATCTTTGACAGGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTG 420
QY 421 GAACATCTTATATCCAAATGTGTCAATACCAACATGGCACATAAACTCGTATCTTAAA 480
DB 421 GAACATCTTATATCCAAATGTGTCAATACCAACATGGCACATAAACTCGTATCTTAAA 480
QY 481 CCTATGTTGCAGAGAGATTTCAATACCTGTGCGACCTCGGGATTTGGATCATATCGCT 540
DB 481 CCTATGTTGCAGAGAGATTTCAATACCTGTGCGACCTCGGGATTTGGATCATATCGCT 540
QY 541 GAGAACATCTGTCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACCTTGTGTGCAAG 600
DB 541 GAGAACATCTGTCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACCTTGTGTGCAAG 600

QY 601 GAATGGTACCGAGTGACCTCTGATGGCATCTCTGGAGAGAGCTTATCAGAGAAATGGTC 660
DB 601 GAATGGTACCGAGTGACCTCTGATGGCATCTCTGGAGAGAGCTTATCAGAGAAATGGTC 660
QY 661 AGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGATGGGAGCAGTATTTATTC 720
DB 661 AGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGATGGGAGCAGTATTTATTC 720
QY 721 AAAAAACAACTCTCTGACGGGAATGCTCTCCCAACTCTCTTTTATAGACACTTTATCT 780
DB 721 AAAAAACAACTCTCTGACGGGAATGCTCTCCCAACTCTCTTTTATAGACACTTTATCT 780
QY 781 AAAATTTATACAGACATTGAGACAAATAGAAATCTAAATGGAGATGTGGAGACATAGTTTA 840
DB 781 AAAATTTATACAGACATTGAGACAAATAGAAATCTAAATGGAGATGTGGAGACATAGTTTA 840
QY 841 CAGAAATTCATCTGCCGAAGTGAAACAAAGAAAGGATTTACGTGTTA CAGTATGATGAT 900
DB 841 CAGAAATTCATCTGCCGAAGTGAAACAAAGAAAGGATTTACTGTTTACAGTATGATGAT 900
QY 901 CAGAAATAGTAAAGCGGCTTTCAGAGACAAACAAATCAAGATCTGGGATAAAACACATTG 960
DB 901 CAGAAATAGTAAAGCGGCTTTCAGAGACAAACAAATCAAGATCTGGGATAAAACACATTG 960
QY 961 GAATGCAAGGAAATTTCTCAGAGCCCATACAGGTTCACTCTCTCTCCAGTATGATGAG 1020
DB 961 GAATGCAAGGAAATTTCTCAGAGCCCATACAGGTTCACTCTCTCTCCAGTATGATGAG 1020
QY 1021 AGAGTGATCATACAGAGATCATCGGATTCACAGGTCAGAGTGGGATCTAAATACAGGT 1080
DB 1021 AGAGTGATCATACAGAGATCATCGGATTCACAGGTCAGAGTGGGATCTAAATACAGGT 1080
QY 1081 GAAATGCTAAACACAGTTGATTCACCAATGTGAAGCAGTTCTGCACTTGGCTTCAATAAT 1140
DB 1081 GAAATGCTAAACACAGTTGATTCACCAATGTGAAGCAGTTCTGCACTTGGCTTCAATAAT 1140
QY 1141 GGCATGATGTGACCTGCTCCAAAGATCGTTCCAAAGATCGTTCCATTTGATGGGATATGCCCTCCCA 1200
DB 1141 GGCATGATGTGACCTGCTCCAAAGATCGTTCCAAAGATCGTTCCATTTGATGGGATATGCCCTCCCA 1200
QY 1201 ACTCAGACTTACCTCTCGGAGGGTGTGGTCCGACACCGAGCTGCTCTCAATGTTGTAGAC 1260
DB 1201 ACTCAGACTTACCTCTCGGAGGGTGTGGTCCGACACCGAGCTGCTCTCAATGTTGTAGAC 1260
QY 1261 TTTGATGACAGTACATTTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA 1320
DB 1261 TTTGATGACAGTACATTTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA 1320
QY 1321 AGTACTTGTGAATTTGTGAAGGACCTTAAATGGACACAAACGAGGCAATGCCCTGTTGCGAG 1380
DB 1321 AGTACTTGTGAATTTGTGAAGGACCTTAAATGGACACAAACGAGGCAATGCCCTGTTGCGAG 1380
QY 1381 TACAGGACAGGCTGTGTGAGTGGCTCATCTGACAAACACTATCAGATTTATGGGACATA 1440
DB 1381 TACAGGACAGGCTGTGTGAGTGGCTCATCTGACAAACACTATCAGATTTATGGGACATA 1440
QY 1441 GAATGTGTCATGTTTACGAGTGTAGAGGCGCATGAGGAATGGTGGCTGTGTTTCCA 1500
DB 1441 GAATGTGTCATGTTTACGAGTGTAGAGGCGCATGAGGAATGGTGGCTGTGTTTCCA 1500
QY 1501 TTTGATAAACAGAGGATGATGATGAGGCGCTATGATGAGGAAATTAAGTGTGGGATCTTT 1560
DB 1501 TTTGATAAACAGAGGATGATGATGAGGCGCTATGATGAGGAAATTAAGTGTGGGATCTTT 1560
QY 1561 GTGGCTGTTTGGACCCCGTGTCTGAGGAGACACTGTCTCAGGACCGCTTGTGGAG 1620
DB 1561 GTGGCTGTTTGGACCCCGTGTCTCAGGAGGACACTGTCTCAGGACCGCTTGTGGAG 1620
QY 1621 CATTCGGGAAAGAGTTTTTTCAGATACAGTTTGTGAATTCAGATTTCTCAGTATGTTCAAT 1680
DB 1621 CATTCGGGAAAGAGTTTTTTCAGATACAGTTTGTGAATTCAGATTTCTCAGTATGTTCAAT 1680

```
QY 1681 GATGACAAATCCTCATCTGGGACTTCTTAATGATCCAGTGCCTCCCAAGCTGAACCCCC 1740
Db 1681 GATGACAAATCCTCATCTGGGACTTCTTAATGATCCAGTGCCTCCCAAGCTGAACCCCC 1740
QY 1741 CGTTCCCTTCTCGAAGATACACCTTACATCTCCAGATAAATAACATACATGACCTCAT 1800
Db 1741 CGTTCCCTTCTCGAAGATACACCTTACATCTCCAGATAAATAACATACATGACCTCAT 1800
QY 1801 ACTTGCCGAGGACCCATTAAAGTTGCGGTATTTAAACGTATCTGCAATACAGGATGAGC 1860
Db 1801 ACTTGCCGAGGACCCATTAAAGTTGCGGTATTTAAACGTATCTGCAATACAGGATGAGC 1860
QY 1861 AACACAGATAAATCAAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGCGGCTT 1920
Db 1861 AACACAGATAAATCAAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGCGGCTT 1920
QY 1921 TGAGACTCCTGTGGGACACAGTTGGTCTGCAGTCGGCCAGGACGGTCTACTCAGCACA 1980
Db 1921 TGAGACTCCTGTGGGACACAGTTGGTCTGCAGTCGGCCAGGACGGTCTACTCAGCACA 1980
QY 1981 ACTGACTGCTTCAGTGTCTATCAGAAGATGTCTTATCAATGTGAATGATGGAAC 2040
Db 1981 ACTGACTGCTTCAGTGTCTATCAGAAGATGTCTTATCAATGTGAATGATGGAAC 2040
QY 2041 TTTTAAACCTCCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 2041 TTTTAAACCTCCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
QY 2101 CAGACAAAGGTGACTTATAATATATTTAGTGTGTTTGGCAGAAAAA 2151
Db 2101 CAGACAAAGGTGACTTATAATATATTTAGTGTGTTTGGCAGAAAAA 2151

RESULT 2
AX057166
LOCUS
DEFINITION
Sequence 42 from Patent WO0075184.
ACCESSION
AX057166
VERSION
AX057166.1 GI:12309979
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Zhang, H., Tsvetkov, L.M. and Kondo, T.
Modulation of protein levels in the scf complex
Patent: WO 0075184-A1
YALE UNIVERSITY (US)
FEATURES
Location/Qualifiers
1..2151
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT
628 a 467 c 513 g 543 t
ORIGIN
Query Match 100.0%; Score 2151; DB 6; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGTTGGCTGGCGCTGGCACCAAGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCG 60
Db 1 TCGTTGGCTGGCGCTGGCACCAAGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCG 60
QY 61 TCGGCGATTATGACCCGCGGAGCGGTGCTGCAAGAGAGGCGACTCAAGTTATGAAT 120
Db 61 TCGGCGATTATGACCCGCGGAGCGGTGCTGCAAGAGAGGCGACTCAAGTTATGAAT 120
QY 121 TCCTCAGAGAGAGAGACTGTATATGCGCAACCCCTAGCAAGATAATACCGAGAG 180
Db 121 TCCTCAGAGAGAGAGACTGTATATGCGCAACCCCTAGCAAGATAATACCGAGAG 180
```

```
QY 181 AATTCACTTAGACAGACATACACAGCTGTGCCAGACTCTCTTAAACCAAGAACAGTA 240
Db 181 AATTCACTTAGACAGACATACACAGCTGTGCCAGACTCTCTTAAACCAAGAACAGTA 240
QY 241 TGTTTAGCAAGCACTGTATGAAGACTGAGAAATTTGTGGCCAAACAAACTTGGCCAAT 300
Db 241 TGTTTAGCAAGCACTGTATGAAGACTGAGAAATTTGTGGCCAAACAAACTTGGCCAAT 300
QY 301 GGCACCTTCAGTATGATTTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAA 360
Db 301 GGCACCTTCAGTATGATTTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAA 360
QY 361 AAGGAACCTGTGTGTCAAATACTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTG 420
Db 361 AAGGAACCTGTGTGTCAAATACTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTG 420
QY 421 GAACATCTTATATCCAAATGTCTATTACCAACATGGGCACATAAACTCGTATCTTAAA 480
Db 421 GAACATCTTATATCCAAATGTCTATTACCAACATGGGCACATAAACTCGTATCTTAAA 480
QY 481 CCTATGTTGCAGAGAGATTTTCATTAATCTCTCTGCGCAGCTCGGGGATTTGGATCATATCGCT 540
Db 481 CCTATGTTGCAGAGAGATTTTCATTAATCTCTCTGCGCAGCTCGGGGATTTGGATCATATCGCT 540
QY 541 GAGAACATTTCTGTCTATCTGATGCCAAATCACTATGTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 GAGAACATTTCTGTCTATCTGATGCCAAATCACTATGTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 GAATGTTACCGAGTGCACCTCTGATGGCATGCTGTGGAAAGAGCTTATCGAGAGAAATGCTC 660
Db 601 GAATGTTACCGAGTGCACCTCTGATGGCATGCTGTGGAAAGAGCTTATCGAGAGAAATGCTC 660
QY 661 AGGACAGATTTCTGTGGAGAGGCTTGCAGAACGAGAGATGGGACAGTATTATTC 720
Db 661 AGGACAGATTTCTGTGGAGAGGCTTGCAGAACGAGAGATGGGACAGTATTATTC 720
QY 721 AAAAAACAACTCTCTGACGGGAATGCTCTCTCCCACTCTTTTATAGAGCACTTTATCCT 780
Db 721 AAAAAACAACTCTCTGACGGGAATGCTCTCTCCCACTCTTTTATAGAGCACTTTATCCT 780
QY 781 AAAATTTATCAAGACATTTGAGACAAATAGAAATCTAATTTGAGATGTGGAAGACATAGTTTA 840
Db 781 AAAATTTATCAAGACATTTGAGACAAATAGAAATCTAATTTGAGATGTGGAAGACATAGTTTA 840
QY 841 CAGAGAAATTCAGTCCGAGAGTGAACAAAGAGGTTTACTGTTTACAGTATGATGAT 900
Db 841 CAGAGAAATTCAGTCCGAGAGTGAACAAAGAGGTTTACTGTTTACAGTATGATGAT 900
QY 901 CAGAAATAATAGTAAGCGGCTTCGAGACAAACAATCAAGATCTGGGATTAACCAACATG 960
Db 901 CAGAAATAATAGTAAGCGGCTTCGAGACAAACAATCAAGATCTGGGATTAACCAACATG 960
QY 961 GAATGCAAGCAATTTCTCAGAGGCAATACAGGTTTCAGTCTCTGCTCCAGTATGATGAG 1020
Db 961 GAATGCAAGCAATTTCTCAGAGGCAATACAGGTTTCAGTCTCTGCTCCAGTATGATGAG 1020
QY 1021 AGAGTGATCATAAACAGGATCATCGGATTCACCGGTTCAGAGTGGGATGTAATACAGGT 1080
Db 1021 AGAGTGATCATAAACAGGATCATCGGATTCACCGGTTCAGAGTGGGATGTAATACAGGT 1080
QY 1081 GAAATGCTAAAACACGTTGATTCACCAATTTGTGAAGCAGTTTCTGCACTTCGGTTTCAATAT 1140
Db 1081 GAAATGCTAAAACACGTTGATTCACCAATTTGTGAAGCAGTTTCTGCACTTCGGTTTCAATAT 1140
QY 1141 GGCATGATGTCACCTGCTCCAAAGATGTTCCATGCTGTATGGGATATGSCCTCCCA 1200
Db 1141 GGCATGATGTCACCTGCTCCAAAGATGTTCCATGCTGTATGGGATATGSCCTCCCA 1200
QY 1201 ACTGACATTTACCTCCGAGGAGGTGCTGTCGACACCGAGCTGCTGCTCAATGTTGTAGAC 1260
Db 1201 ACTGACATTTACCTCCGAGGAGGTGCTGTCGACACCGAGCTGCTGCTCAATGTTGTAGAC 1260
QY 1261 TTTGATGACACAGTACATTTTCTGATCTGGGATAGAACTATATAAGGATAGGAACACA 1320
```


Db	541	GAGAACATTCTGTCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTGCAAG	600
Qy	601	GAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAGAAAGCTTATPCGAGAGAAATGGTC	660
Db	601	GAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAGAAAGCTTATPCGAGAGAAATGGTC	660
Qy	661	AGGACAGATCTCTGTGGAGAGGCTCGCAGAACGAGAGAGATGGGACAGTATTTATTC	720
Db	661	AGGACAGATCTCTGTGGAGAGGCTCGCAGAACGAGAGAGATGGGACAGTATTTATTC	720
Qy	721	AAAAAACAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATPCCT	780
Db	721	AAAAAACAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATPCCT	780
Qy	781	AAATATTATACAGACATTGACACAAATAGAAATCTAAATGGAGATGTGGAAGACATAGTTTA	840
Db	781	AAATATTATACAGACATTGACACAAATAGAAATCTAAATGGAGATGTGGAAGACATAGTTTA	840
Qy	841	CAGAGATTCTACTCCGGAAGTGAACAGCAAGCAAGGCTTTACTCTTTACAGTATGATGAT	900
Db	841	CAGAGATTCTACTCCGGAAGTGAACAGCAAGCAAGGCTTTACTCTTTACAGTATGATGAT	900
Qy	901	CAGAAAAATAGTAAGCGCTTCGAGACAAACAAATCAAGATCTGGGATAAAAAACACATTG	960
Db	901	CAGAAAAATAGTAAGCGCTTCGAGACAAACAAATCAAGATCTGGGATAAAAAACACATTG	960
Qy	961	GAATGCAAGCAATTCACAGGCCATACAGGTTCACTCTCTGCTCCAGTATGATGAG	1020
Db	961	GAATGCAAGCAATTCACAGGCCATACAGGTTCACTCTCTGCTCCAGTATGATGAG	1020
Qy	1021	AGAGTGATCAACAGAGATCATCGGATTCACAGGTCAGAGTGTGGAGTGAATAACAGGT	1080
Db	1021	AGAGTGATCAACAGAGATCATCGGATTCACAGGTCAGAGTGTGGAGTGAATAACAGGT	1080
Qy	1081	GAATGCTAAACACAGTTGATTACCACTTTGAAACAGTCTTGCACTTGCCTTTCAATAAT	1140
Db	1081	GAATGCTAAACACAGTTGATTACCACTTTGAAACAGTCTTGCACTTGCCTTTCAATAAT	1140
Qy	1141	GGCATGATGGTACCTGCTCCAAAGATCGTTCCATCTGCTGATGGATATGGCTCCCA	1200
Db	1141	GGCATGATGGTACCTGCTCCAAAGATCGTTCCATCTGCTGATGGATATGGCTCCCA	1200
Qy	1201	ACTGACATTACCTCCGAGGGTCTGCTGGACACCGAGTCTGCTCAATGTGTAGAC	1260
Db	1201	ACTGACATTACCTCCGAGGGTCTGCTGGACACCGAGTCTGCTCAATGTGTAGAC	1260
Qy	1261	TTTGATGACAGTACATTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA	1320
Db	1261	TTTGATGACAGTACATTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA	1320
Qy	1321	AGTACTTGTAATTTGTAAGACCTTAAATGGACACAAACGAGGCATTCCTGTTCCAG	1380
Db	1321	AGTACTTGTAATTTGTAAGACCTTAAATGGACACAAACGAGGCATTCCTGTTCCAG	1380
Qy	1381	TACAGGACAGGCTGTAGTGAAGGCTTCATCTGACAAACACTATCAGATTATGGACATA	1440
Db	1381	TACAGGACAGGCTGTAGTGAAGGCTTCATCTGACAAACACTATCAGATTATGGACATA	1440
Qy	1441	GAATGTGTGATGTTTACGAGTGTGAAGGCATGAGGAATTTGGTGGCTGTGATTCGA	1500
Db	1441	GAATGTGTGATGTTTACGAGTGTGAAGGCATGAGGAATTTGGTGGCTGTGATTCGA	1500
Qy	1501	TTTGATAACAGAGATAGTCAAGTGGGCTTATGATGGAAAAATTAAGGTGGGATCTT	1560
Db	1501	TTTGATAACAGAGATAGTCAAGTGGGCTTATGATGGAAAAATTAAGGTGGGATCTT	1560
Qy	1561	GTGGTGTCTTTGGACCCCGTCTCTCGAGGACACTCTGCTACGACCCCTTTGTGGAG	1620
Db	1561	GTGGTGTCTTTGGACCCCGTCTCTCGAGGACACTCTGCTACGACCCCTTTGTGGAG	1620
Qy	1621	CATTCCGGAAGATTTTTCGACTACAGTTTGAATTCAGATTCAGATTCAGTTCACAT	1680
Db	1621	CATTCCGGAAGATTTTTCGACTACAGTTTGAATTCAGATTCAGTTCACAT	1680

Qy	1681	GATGACAAATCCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCCC	1740
Db	1681	GATGACAAATCCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCCC	1740
Qy	1741	CGTTCCCTTCTCGAAACATACACCTACATCTCCAGATAAATAACCATACACTGACCTCAT	1800
Db	1741	CGTTCCCTTCTCGAAACATACACCTACATCTCCAGATAAATAACCATACACTGACCTCAT	1800
Qy	1801	ACTTCCCGCAGGACCCATTAAAGTTCGGGTATTTAACTATCTGCAATACCCAGGATGAGC	1860
Db	1801	ACTTCCCGCAGGACCCATTAAAGTTCGGGTATTTAACTATCTGCAATACCCAGGATGAGC	1860
Qy	1861	AACAACAGTAAACAATCAAACTACTGCCCAGTTTCCCTGGACTAGCCGAGGAGCAGGGCTT	1920
Db	1861	AACAACAGTAAACAATCAAACTACTGCCCAGTTTCCCTGGACTAGCCGAGGAGCAGGGCTT	1920
Qy	1921	TCGAGACTCTGTGTTGGACACAGTTGCTGCGATCGGCCCGCAGGAGCGTCTACTCAGCACA	1980
Db	1921	TCGAGACTCTGTGTTGGACACAGTTGCTGCGATCGGCCCGCAGGAGCGTCTACTCAGCACA	1980
Qy	1981	ACTGACTGCTTCAGTCTCTATCAGAAAGATGTCTTCTATCAATTTGGAATGATGGAAC	2040
Db	1981	ACTGACTGCTTCAGTCTCTATCAGAAAGATGTCTTCTATCAATTTGGAATGATGGAAC	2040
Qy	2041	TTTTAAACCTCCCTCTCTCTCTTTCACCTCTGACCTAGTTTTCCTCATTTGGTTC	2100
Db	2041	TTTTAAACCTCCCTCTCTCTCTTTCACCTCTGACCTAGTTTTCCTCATTTGGTTC	2100
Qy	2101	CAGACAAAGGTGACTTATAATATATTTAGTGTGTTTGGCAGAAAAAAA	2151
Db	2101	CAGACAAAGGTGACTTATAATATATTTAGTGTGTTTGGCAGAAAAAAA	2151

RESULT 4	AF129530	2151 bp	mRNA	linear	PRI 31-OCT-1999
LOCUS	Homo sapiens chromosome 10 F-box protein FBW1A (FBW1A) mRNA,				
DEFINITION	complete cds.				
ACCESSION	AF129530				
VERSION	AF129530.1	GI:6164609			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Cenciarelli, C., Chiar, D.S., Guardavaccaro, D., Parks, W., Vidal, M. and Pagano, M.				
TITLE	Identification of a family of human F-box proteins				
JOURNAL	Journal of Molecular Biology	9 (20), 1177-1179			(1999)
MEDLINE	10531035				
PUBMED	10531035				
REFERENCE	2 (bases 1 to 2151)				
AUTHORS	Chiar, D.S. and Pagano, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-FEB-1999) Pathology, NYU Medical Center, 550 First Ave. MSB 548, New York, NY 10016, USA				
FEATURES	Location/Qualifiers				
source	1. .2151				
gene	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="10"				
	/map="10q24"				
	1. .2151				
	/gene="FBW1A"				
	70. .1779				
	/gene="FBW1A"				
	/function="regulates the stability of IkappaB alpha- and beta-catenin"				
	/note="contains one F-box domain and seven WD40 domains; forms an SCF with Skp1 and Cull1"				
CDS					

Db 1222 ACTGACATCACCTCAGAGGGTCTCTGGTGGACACCGAGCTCGGTCATATGTTGTAGAC 1281
 Qy 1261 TTTGATGACAAAGTACATGTTCTGTCATCTGGGATAGAACTATAAGGTATGGAACACA 1320
 Db 1282 TTTGATGACAAAGTACATGTTCTGTCATCTGGGATAGAACTATAAGGTATGGAACACA 1341
 Qy 1321 AGTACTGTGTAATTTGTAAGACCTTAAATGGACACAAACGAGGCAATGCTGTTGGAG 1380
 Db 1342 AGTACTGTGTAATTTGTAAGACCTTAAATGGGACAAAGGCTGGCATCGCTGTTGGAG 1401
 Qy 1381 TACAGGACAGGCTGTAGTGGAGGCTCATCTGACACACACTATCAGATTTATGGGACATA 1440
 Db 1402 TACAGACAGGCTGTGGTGGAGGCTCTCTGACACACACTATCAGCTGTGGGACATA 1461
 Qy 1441 GAATGTGTGCATGTTTACGAGTGTTAAGAGGCGCATGAGGAATGTTGCGTGTATTGGA 1500
 Db 1462 GAGTGTGGAGCATGCTCGAGTGTGAGGGCGCATGAGGATGTTGTCGCTGATTCGA 1521
 Qy 1501 TTTGATACAGAGGATAGTACGAGGCTCATGATGGAATAATTAAGTGTGGATCTT 1560
 Db 1522 TTTGATACAAAGGATAGTACGAGGCTCATGATGGAATAATTAAGTGTGGATCTT 1581
 Qy 1561 GTGGCTGTTTGGACCCCGCTCTCTGCGAGGACACTCTGTCTACGACCTTGTGGAG 1620
 Db 1582 ATGGCTGTTTGGACCCCGCTCTCTGCGAGGACACTCTGTCTGCGGACACTTGTGGAG 1641
 Qy 1621 CATTCGGAAGAGTGTTCGATACAGTTGATGAATTCAGATTTGTCAGTAGTTCACAT 1680
 Db 1642 CATTCGGAAGAGTGTTCGCGCTCCAGTTTGTATGAATTCAGATTTGTCAGTAGTTCACAT 1701
 Qy 1681 GATGACAAATCTCATCTGGGATCTCTAAATGATCCAGTCCGACGCTGAACCCGCC 1740
 Db 1702 GATGACAAATCTCATCTGGGATCTCTGATGATCCAGTCTCACTGCTGAACCGGCC 1761
 Qy 1741 CGTTCCTCTCTGGAACATACACCTATCTCCAGATAAATAACATACATACACCTCAT 1800
 Db 1762 CGCTCCCTCTCTGGAACATACACCTATCTCCAGATAAATAACATACATACACCTCAT 1821
 Qy 1801 ACTTGCCAGACCCATTAAGTTCGCTATTTAACTGCTATCTGCAATACAGGATGAGC 1860
 Db 1822 AATTGCCAGGATTCGTTAATGTTGCGATTTTAAACAGCTGCAAGACAGGATGAAC 1881
 Qy 1861 AACACAGTAACAACTACTGCCAGTTCCTCCCTG--GACTAGCCGAGGACGAGGC 1918
 Db 1882 -----AACATCAACTCTTACCGGATTCGCGAGGATGAGCGAGGACGAGGC 1932
 Qy 1919 TTTGAGACTCTGTTGGACACAGTTGTTGCGAGTTCGCGAGGACGCTTACTAGCA 1978
 Db 1933 TTTGAGACTCTGTTGGACACAGTTCGCTCAGCAG--CCGACGAGGACGCTCTCGGC- 1990
 Qy 1979 CAACCTGCTGCTGCTATCAGAGATGCTCTTATCAATTTGGAATGATGGA 2038
 Db 1991 --ACCGCTGCTGCTGCTATCAGAGATGCTCTT-TATCTTGTGATGATGGA 2047
 Qy 2039 ACTTTTAAAGCTCCCTCTCTCC---TCCTTTCACTCTGACCTAGTCTTTTCCCAT 2095
 Db 2048 ACTTCAAGCTCCCTCTCTCCCTTCCCTTCCCTTCCCTGACCTGCTTCCCTCCCAT 2107
 Qy 2096 -GGTTCACAGAAAGGCTGATTAATAATATTAGTGTGTTTGGCAGAAAAA 2151
 Db 2108 GGGTTCACAGAAAGGCTGATTAATAATATTAGTGTGTTTGGCAGAAAAA 2164

RESULT 8
 AF101784
 LOCUS AF101784 1818 bp mRNA linear PRI 21-JAN-1999
 DEFINITION Homo sapiens b-TRCP variant E3RS-IkappaB mRNA, partial cds.
 ACCESSION AF101784
 VERSION AF101784.1 GI:4165135
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 1818)
 Yaron,A., Hatzubai,A., Davis,M., Lavan,I., Amit,S., Manning,A.M.,
 Andersen,J.S., Mann,M., Mercurio,F. and Ben-Neriah,Y.
 Identification of the receptor component of the
 IkappaBalpha-ubiquitin ligase
 Nature 396 (1998), 590-594 (1998)
 99075339
 9859996
 2 (bases 1 to 1818)
 Yaron,A., Hatzubai,A., Mercurio,F., Mannig,A.M., Andersen,J.S.,
 Mann,M. and Ben-Neriah,Y.
 Direct Submission
 Submitted (27-OCT-1998) Immunology, Hebrew University-Hadassah
 Medical School, Jerusalem, Israel
 Location/Qualifiers
 1. 1818
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 <1. 1818
 /note="beta-transducin repeats-containing protein;
 receptor component of IkappaBalpha-ubiquitin ligase; contains
 F-box and WD regions"
 /codon_start=1
 /product="b-TRCP variant E3RS-IkappaB"
 /protein_id="AA008702.1"
 /db_xref="GI:4165136"
 /translation="MDPAEAVLQEKALFKWCSMPRLSLGCSLADSPSLRLVNP
 TGLATAPONSEREDCNPPPKIIPKNSLQTYNSCARCLNGETVCLASTAMKT
 ENCVAKTKLANGTSSMIVKORSLKSYKEKELCKVKEFQMSQVDFVHLISOM
 CHYGHGINSYLVKPLQDFTALPARGLDHAENILSYLDAKSLCAELVCKEWRV
 TSDGMJWKLLIEMVSTDSLRGLAERGGGOVLKPKPDGNAPNSFVRLVYKII
 QDLTIESWRCRHSRLQHCRESYSGVYCLQYDDQKIVSLRNTIKINDKNTLE
 CKRLTGHVSILQIDERVITSSDSTVRVWNTVMGMLNTLLHHCVAVLHURV
 NGMWTGCKDRIAVMDMSPTDITLRLVLRVRAAVNVVDFDKYIVSAGSDRTIKV
 WNTCTCFEVLNKGKGIACQVDRDLVVGSSDNTIRLWDIECCAGLRVLEHLEL
 VRCIRPFNKIRVSGVAGDKIKVMDLVAALDPRAPAGTLCRLTLVHSGRVFRLQDFEF
 QIVSSSHDDILLIWDPLNDPAOAEPRSPRTYTVISR"
 532 a 385 c 445 g 456 t
 BASE COUNT
 ORIGIN
 Query Match 77.3%; Score 1663.2; DB 9; Length 1818;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1665; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 112 TTTATGATTTCTCAGAGAGAGAGAGCTGTAATAATGGCGAACCCCTAGGAGAGATAA 171
 Db 151 TTCCAGAAATTCCTCAGAGAGAGAGAGCTGTAATAATGGCGAACCCCTAGGAGAGATAA 210
 Qy 172 CCAGAGAGAAATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTTAAACCAA 231
 Db 211 CCAGAGAGAAATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTTAAACCAA 270
 Qy 232 GAACAGTAGTGTAGACAGACTGCTATGAAAGCTGAGAAATGTGTGGCCAAACAAA 291
 Db 271 GAACAGTAGTGTAGACAGACTGCTATGAAAGCTGAGAAATGTGTGGCCAAACAAA 330
 Qy 292 CTTCGCAATGCACTTCCTCAGTATGATTGTGCCCAAGCAACGAAATCTTCAGCAAGCTAT 351
 Db 331 CTTCGCAATGCACTTCCTCAGTATGATTGTGCCCAAGCAACGAAATCTTCAGCAAGCTAT 390
 Qy 352 GAAAGGAAAGAGAACTGTGTCAAATCTTTTACAGAGTGGTTCAGAGTTCAGATCAAGTG 411
 Db 391 GAAAGGAAAGAGAACTGTGTCAAATCTTTTACAGAGTGGTTCAGAGTTCAGATCAAGTG 450
 Qy 412 GAATTTGTGGAACATCTTATATCCCAATGTGTCAATACCAATGGGACATATAACTCG 471
 Db 451 GAATTTGTGGAACATCTTATATCCCAATGTGTCAATACCAATGGGACATATAACTCG 510
 Qy 472 TATCTTAAACCTATGTTGACAGAGATTTTCATTAACCTGCTCTGCCAGCTGGGATTTGGAT 531
 Db 511 TATCTTAAACCTATGTTGACAGAGATTTTCATTAACCTGCTCTGCCAGCTGGGATTTGGAT 570

QY	532	CATATCGCTGAGAACATTTCTGTCTATACCTGGATGCGCAATCACTATGTGCTGCTGAACTT	591
Db	571	CATATCGCTGAGAACATTTCTGTCTATACCTGGATGCGCAATCACTATGTGCTGCTGAACTT	630
QY	592	GTGTGCAAGGAATGCTACCGAGTGACCTCTGATGCGCATGCTGTGGAGAGCTTATCGAG	651
Db	631	GTGTGCAAGGAATGCTACCGAGTGACCTCTGATGCGCATGCTGTGGAGAGCTTATCGAG	690
QY	652	AGAAATGCTCAGGACAGATTTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAG	711
Db	691	AGAAATGCTCAGGACAGATTTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAG	750
QY	712	TATTTATTCACAAACAACTCTGACGGGAATGCTCTCCCACTCTTTTATAGAGCA	771
Db	751	TATTTATTCACAAACAACTCTGACGGGAATGCTCTCCCACTCTTTTATAGAGCA	810
QY	772	CTTTATCTCTAAATATATACAGACATTTGAGACAAATAGAACTTAATTTGGAGATGTGGAAGA	831
Db	811	CTTTATCTCTAAATATATACAGACATTTGAGACAAATAGAACTTAATTTGGAGATGTGGAAGA	870
QY	832	CATAGTTTACAGAGATTTCACTCCGAACTGAAACAGCAAGAGGTTTACITGTTTACAG	891
Db	871	CATAGTTTACAGAGATTTCACTCCGAACTGAAACAGCAAGAGGTTTACITGTTTACAG	930
QY	892	TATGATGATCAGAAAAATAGTAAGCGGCTTTCGAGACAAACAAATCAAGATCTGGGATAAA	951
Db	931	TATGATGATCAGAAAAATAGTAAGCGGCTTTCGAGACAAACAAATCAAGATCTGGGATAAA	990
QY	952	AACACATGGAATGCAAGCGAATTTCTCAGCGCCATACAGGTTGAGTCTCTGCTCCAG	1011
Db	991	AACACATGGAATGCAAGCGAATTTCTCAGCGCCATACAGGTTGAGTCTCTGCTCCAG	1050
QY	1012	TATGATGAGAGATGATCATAACAGGATCATCGGATTCACGGTCAGAGTGTGGAGTGA	1071
Db	1051	TATGATGAGAGATGATCATAACAGGATCATCGGATTCACGGTCAGAGTGTGGAGTGA	1110
QY	1072	AATACAGGTGAATGCTAACACGTTGATTCACCATGTTGAGAGGTTCTGCATTTGCGT	1131
Db	1111	AATACAGGTGAATGCTAACACGTTGATTCACCATGTTGAGAGGTTCTGCATTTGCGT	1170
QY	1132	TTCAATTAATGGCATGATGTGACCTGCTCCAAAGATCGTTCCATTCGTTGATGGATATG	1191
Db	1171	TTCAATTAATGGCATGATGTGACCTGCTCCAAAGATCGTTCCATTCGTTGATGGATATG	1230
QY	1192	GCCTCCCAACTGACATTAACCTCCGAGGGTCTGTCGACACCGAGCTGCTGCAAT	1251
Db	1231	GCCTCCCAACTGACATTAACCTCCGAGGGTCTGTCGACACCGAGCTGCTGCAAT	1290
QY	1252	GTGTAGACTTTGATGACAGTACATTTGTTCTGTCATCTGGGGATAGAACTATAAGGTA	1311
Db	1291	GTGTAGACTTTGATGACAGTACATTTGTTCTGTCATCTGGGGATAGAACTATAAGGTA	1350
QY	1312	TGGAACACAGTACTTGTGAATTTGAAGGACTTAATTTGACACAAACGAGGATGACC	1371
Db	1351	TGGAACACAGTACTTGTGAATTTGAAGGACTTAATTTGACACAAACGAGGATGACC	1410
QY	1372	TGTTTCAGTACAGGACAGGCTGGTAGTGAGTGGCTCATCTGCAACACATATCAGATTA	1431
Db	1411	TGTTTCAGTACAGGACAGGCTGGTAGTGAGTGGCTCATCTGCAACACATATCAGATTA	1470
QY	1432	TGGGACATAGAAATGTTGGTGTGATTTTACAGTGTGTTAGAGGCCATGAGGAATTTGTCGT	1491
Db	1471	TGGGACATAGAAATGTTGGTGTGATTTTACAGTGTGTTAGAGGCCATGAGGAATTTGTCGT	1530
QY	1492	TGTATTCGATTTGATTAACAGAGGATAGTCACTGGGGCTTATGATGGAAAAATTAAGTG	1551
Db	1531	TGTATTCGATTTGATTAACAGAGGATAGTCACTGGGGCTTATGATGGAAAAATTAAGTG	1590
QY	1552	TGGGATCTTTGGCTGCTTTGGACCCCGTGTCTCTGAGGACACTCTGCTCTACGGACC	1611
Db	1591	TGGGATCTTTGGCTGCTTTGGACCCCGTGTCTCTGAGGACACTCTGCTCTACGGACC	1650

QY	1612	CTTTGGAGCAATTCGGAGAGTTTTTCGACTACATTTTGATGAATTCAGATTTGTCAGT	1671
Db	1651	CTTTGGAGCAATTCGGAGAGTTTTTCGACTACATTTTGATGAATTCAGATTTGTCAGT	1710
QY	1672	AGTTCCATGATGACACAATCCTCANTCGGACTTCTTAATGATCCAGCTGCCAAGCT	1731
Db	1711	AGTTCCATGATGACACAATCCTCANTCGGACTTCTTAATGATCCAGCTGCCAAGCT	1770
QY	1732	GAACCCCGGCTTCCCTTTCGAAACATACCATCATCTCCAGATAA	1779
Db	1771	GAACCCCGGCTTCCCTTTCGAAACATACCATCATCTCCAGATAA	1818

BC003989	2081 bp	mRNA	linear	ROD 16-APR-2003
Mus musculus beta-transducin repeat containing protein, mRNA (cdna clone MGC:7517 IMAGE:3491843), complete cds.				
BC003989				
BC003989.1	GI:13278339			
MGC:				
Mus musculus (house mouse)				
Mus musculus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
1 (bases 1 to 2081)				
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schreitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Mullay, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.				
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
Proc. Natl. Acad. Sci. U.S.A.	99 (26)	16899-16903	(2002)	
2388257				
12477932				
2 (bases 1 to 2081)				
Strausberg, R.				
Direct Submission				
Submitted (28-FEB-2001)				
Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
NIH-MGC Project URL: http://mgc.nci.nih.gov				
Contact: MGC help desk				
Email: cgapbs-remail.nih.gov				
Tissue Procurement: Gilbert Smith, Ph.D.				
CDNA Library Preparation: Life Technologies, Inc.				
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)				
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305				
Web site: http://www-shgc.stanford.edu				
Contact: (Dickson, Mark) mdcpaxil.stanford.edu				
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAC Plate: 8 Row: 0 Column: 21
 This clone was selected for full length sequencing because it


```

Db 901 CGGATTCCTCAGGCGCCACACAGGGCTCGTCTGTCTGCGAGTACGATGAGAGGCTGATC 960
QY 1030 ATACAGGATCATCGGATTCACGCTCAGAGTGGGATGTAATAATACAGTCAATGCTA 1089
Db 961 ATACAGGCTCTCAGACTCCACGCTCAGAGTGGGATGTAATAATGAGTGAATGCTA 1020
QY 1090 AACAGGTGATTCACCATTTGAGCAGATTTCTCAGCTTGGTTTCAATATGCGATG 1149
Db 1021 GACACATGACTCACCACCTGTAAGCGGTTCTGCACCTGGCTTCAATATGCGATG 1080
QY 1150 GTGACCTCTCCAAAGATCGTTCCATTGCTGATGGATATGCCCTCCCACTGACATT 1209
Db 1081 GTGACCTGTTTCCAAAGACCGTTCCATGCTGTGGGATATGGCTTCCCACTGACATC 1140
QY 1210 ACCCTCGAGGGTGTGTCGGACACCGAGCTGCTGTCATATGTTGTAGACTTTGATGAC 1269
Db 1141 ACCCTCAGAGGGTGTGTCGGACACCGAGCTGCGTCAATGTTGTAGACTTTGATGAC 1200
QY 1270 AAGTACATTTGTTCTGCTATCGGGATAGAACTATTAAGGTATGGAACCAAGTCTGT 1329
Db 1201 AAGTACATCGTTTCTGCTCTCGGCTGAGATAGAACCATTAAGGTGTGGAACCAAGTCTGT 1260
QY 1330 GAATTTCTAAGGACCTTAAATGGACACACAAACGAGGCAATGCTGTTTGGAGTACAGGAC 1389
Db 1261 GAATTCGTAAGGACCTTAAATGGGACACAGCGTGGCATCGCTGTTTGGAGTACAGGAC 1320
QY 1390 AGGCTGTAGTGTGCTCATCTGACAACTATCAGATTTATGGACATAGAAATGTGT 1449
Db 1321 AGGCTGTGTGTGAGCGCTCTCTGACACACCATCAGGCTGTGGACATAGAGTGTGA 1380
QY 1450 GCATGTTTACAGTGTGTAGAGGACCATGAGGAATGCTGTTGATTCGATTTGATAC 1509
Db 1381 GCATGCTGTGAGTGTGTGAGGACCATGAGAGTGTGTACGCTGCATTCGATTTGATAC 1440
QY 1510 AAGAGGATGTAGTGGGCTATGATGGAAAAATTAAGTGTGGGATCTTGTGGCTGT 1569
Db 1441 AAGAGGATGTAGGCGGACCTATGATGGGAAAAATTAAGTGTGGGATCTTATGGCTGT 1500
QY 1570 TTGACCCCGTGTCTCGCAGGACACTGTGCTAGGACCGTGTGGAGATTCGGA 1629
Db 1501 TTGACCCCGTGTCTCGCAGGAGACTCTCTGTCTGGACACTTGTGGAGATTCGGA 1560
QY 1630 AGAGTGTTCGACTACATGTTGATGAATTCAGATTTGATGATGTTTCAATGATGACACA 1689
Db 1561 AGAGTGTTCGCTCCAGTTTGTATGAATTCAGATTTGATGATGTTTCAATGATGACACA 1620
QY 1590 ATCTCATCTGGGACTTCTTAATGATCCAGTGTGCGCAAGTGAACCCCGCTCCCT 1749
Db 1621 ATCTCATCTGGGACTTCTTGAATGATCCAGTGTGCTCAGCTGAACCCCGCTCCCT 1680
QY 1750 TCTCGAATACATACACTACATCTCCAGATTAATTAACCATACACTGACCTCATCTTGCCTCA 1809
Db 1681 TCTCGAATACATACACTACATCTCCAGATTAATTAACCATACACTGACCTCATCTTGCCTCA 1740
QY 1810 GGACCCATTAAGTGTGGTATTAAGTATTCGCAATTCGCAATTCAGAGTACGCAACAGT 1869
Db 1741 GGACTCGTTAATGTTGAGTATTAACAGACCTGCGCAAGACAGGATGAAC----- 1791
QY 1870 AACATCAACTACTGCTCCAGTTTCCCTG--GACTAGCGAGGAGCAGGCTTTCAGACT 1927
Db 1792 AACATCAACTCTTACCGGATTCGCGAGGATGAGCGAGGAGCAGGCTTTCAGACT 1851
QY 1928 CCTGTTGGGACACAGTTGGTCTGAGTTCGCGCCAGGACGCTCTACTCAGCACAATGACT 1987
Db 1852 CCTGTTGGGACACAGTCCGTCAGCAG--CCGACCGAGGACGCTCTGCTCCG--ACCGGCT 1907
QY 1988 GCTTCAGTGTCTATCAGAGATGCTTCTTATCAATGTGAATGATGGAATTTTAA 2047
Db 1908 GCCTCAGTGTCTATCAGAAGATGCTT--TATCTTGTGTGAATGATGGAATTTTAA 1966
QY 2048 CCTCCCTCC 2057
Db 1967 CCTCCCTCC 1976

```

```

RESULT 11
AF099932
LOCUS AF099932
DEFINITION Mus musculus beta-TrCP protein E3RS-IkappaB mRNA linear ROD 21-JAN-1999
ACCESSION AF099932
VERSION AF099932.1 GI:4008019
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1712)
Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A. M.,
Andersen, J. S., Mann, M., Mercurio, F. and Ben-Neriah, Y.
Identification of the receptor component of the
IkappaBalpha-ubiquitin ligase
Nature 396 (6711), 590-594 (1998)
99075339
9859996
2 (bases 1 to 1712)
Yaron, A., Hatzubai, A., Mercurio, F., Manning, A. M., Andersen, J. S.,
Mann, M. and Ben-Neriah, Y.
Direct Submission
Submitted (19-Oct-1998) Immunology, Hebrew University of Jerusalem,
Bin Karem, Jerusalem 91120, Israel
Location/Qualifiers
1..1712
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..1710
/notes="beta-transducin repeat-containing protein; receptor
component of IkBa-ubiquitin ligase"
/codon_start=1
/product="beta-TrCP protein E3RS-IkappaB"
/protein_id="A008020.1"
/db_xref="GI:4008020"
/translation="MDPAVLQELKFMNNSEREDCNNGEPKRIIPKNSLRQTY
NSCARLINCNETVLTSTAMTENCVCYAKLANGTSMIVPKRKLASYPKEKELCV
KYPEOWSESDQVEFVEHLTISOMCHYQHGHINSYLPMLQDFITLARGDHLAENI
LSYLDKSLCAAEVLCKEYRVTSDGLMKWKLIERMVTQSLRGLARRGQWGLPK
DKPPDENAPNSFYALPKIIOJLTIETSNWRGCRHSLQHCSETSKGVCLQYD
DOKIUSGLRDNTRIKINDKSTLECKAILGHTGSLCLQYDERVILITSSDSYRMDV
NAGEMLNTLIHCEAVLHLRFNNGMMVTCSDRSIAYVDMASPTDITLRLVVGHRRA
VNVVDFDKIIVSASGDRITIKVWNTSTCEFTVRLNGHKRGIACLOYRDLRVVSGSSN
TIRLWDIEGACLRVLEGEHELVRCIFDNKRIIVSGAYDKIKVMDLMAALDPRAPAG
TLCRLTLVHSGRVRFLQDFEQIVSSSHDDTLITWDFLNDPAAHAEPSPSRITY
ISR"
BASE COUNT 469 a 399 c 453 g 391 t
ORIGIN
Query Match 67.5%; Score 1451.2; DB 10; Length 1712;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 1549; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
QY 70 ATGAGCCCGCGAGCGGTGCTGCAAGAGAGGACCTCAAGTTTATGAAATTCCTCAGAG 129
Db 1 ATGAGCCCGCGAGCGGTGCTGCGAGAGAACGCTTAAGTTTATGAAATTCCTCAGAG 60
QY 130 AGAAGAAGACTGTAATAATGGGAACCCCTAGGAAGATAATACCGAGAAGAATTCACCT 189
Db 61 AGAAGAAGACTGTAATAATGGGAACCCCTAGGAAGATAATACCGAGAAGAATTCACCT 120
QY 190 AGACAGACATACAGAGCTGTGCCAGCTCTCTTAAACCGAAGACAGATGTTTACGA 249
Db 121 AGACAGACTTACACAGCTGTGCCAGCTTTCATATAAACCAAGACAGATGTTGCTAAC 180
QY 250 AGCACTGCTATGAAGACTGAGAAATTTGTGTGGCCAAAAACAAACTTGCCTTCC 309
Db 181 AGCACTGCTATGAGACTGAGAAATTTGTGTGGCCAAAGCCAACTTGCCTTCC 240

```

QY 310 AGTATGATTGTGCGCAACGACGAACTCTCAGCAAGCTATGAAAGGAAAGGAACTG 369
Db |||||
QY 241 AGCATGATTGTGCGCAACGACGCGAACTCTCAGCAAGCTATGAAAGGAAAGGAGCTG 300
Db |||||
QY 370 TGTGTCAAAATATCTTTGAGCAGTGTGAGAGTCAAGTCAAGTGGAAATTTGTGGAACATCTT 429
Db |||||
QY 301 TGTGTCAAGTATTTTGTGAGCAGTGTGAGAGTCTGATCAAGTGGAAATTTGTGGAACACCTT 360
Db |||||
QY 430 ATATCCCAATGTGTCAATCAACATGGGACATAAATCTGATCTTTAAACCTATGTTG 489
Db |||||
QY 361 ATATCCCAATGTGTCAATCAACATGGGACATAAATCTGATCTTTAAACCTATGTTG 420
Db |||||
QY 490 CAGAGAGATTTTCAATACCTGTCTGCCAGCTCGGGATTTGGATCATATCGCTGAGAACAT 549
Db |||||
QY 421 CAGAGGATTTTCAATACCTGTCTGCCAGCTCGGGATTTGGATCATATCGCTGAGAACAT 480
Db |||||
QY 550 CTGTATACCTGGATGCGCAATCACTATGCTGCTGAACTTTGTGTGCAAGGATGGTAC 509
Db |||||
QY 481 CTGTATACCTGGATGCGCAATCACTATGCTGCTGAACTTTGTGTGCAAGGATGGTAC 540
Db |||||
QY 610 CGAGTGACCTCTGTATGCGATGCTGTGGAAGAGCTTATCGAGAGATGGTTCAGGACAGAT 669
Db |||||
QY 541 CGCGTGACCTCGGACGCGCATGCTGTGGAAGAGCTCATCGAGAGATGGTTCAGGACGAC 600
Db |||||
QY 670 TCTCTGTGAGAGGCTCGGACGAGAGAGAGATGGGACAGTATTTATTTCAAAAACAAA 729
Db |||||
QY 601 TCTCTGTGCGAGGCTCGGACGAGAGAGAGCTGGGACAGTATTTATTTCAAAAACAAA 660
Db |||||
QY 730 CTTCTGTGCGGAAATGCTCTCTCCAACTCTTTTATAGACACTTTATCTCTAAATATA 789
Db |||||
QY 661 CTTCTGTGAGAGCTCTCTCCAACTCTTTTATAGAGCTTTATCTCTAAATATA 720
Db |||||
QY 790 CAAAGCATTTGAGCAATAGATCTAATTTGGAGATGTGGAAGACATAGTTTACAGAAAT 849
Db |||||
QY 721 CAAAGCATTTGAGCAATAGATCTAATTTGGAGATGTGGAAGACATAGTTTACAGAAAT 780
Db |||||
QY 850 CACTGCGGAAGTGAACAAAGAGGATTTACTGTTTACAGTATGATGATCAGAAATA 909
Db |||||
QY 781 CACTGCGGAGTGAACAAAGGATTTACTGTTTACAGTATGATGATCAGAAATA 840
Db |||||
QY 910 GTAAGGCGCTTCGAGACAAACAAATCAAGATCTGGGATTAACAAACATTTGGAAATGCAAG 969
Db |||||
QY 841 GTGAGGCGCTTCGAGACAAACAAATCAAGATCTGGGATTAACAAACATTTGGAAATGCAAG 900
Db |||||
QY 970 CGAATTTCTCAGAGGCGATACAGGTTTCACTCTCTGCTCCAGTATGATGAGAGATGATC 1029
Db |||||
QY 901 CGGATTTCTCAGGCGCCACAGGCTCCGCTCTGCTGTGCTGCAATGATGAGAGGATGATC 960
Db |||||
QY 1030 ATACAGGATCATCGGATTCACGGTCAAGTGTGGGATGTAATACAGTGAATGCTTA 1089
Db |||||
QY 961 ATCAGAGGCTCTCAGATCTCCACCGTCAAGTGTGGGATGTAATACAGTGAATGCTTA 1020
Db |||||
QY 1090 AACAGGTTGATTCACCATTTGAAGCAGTCTGCACTTTCGCTTTCAATATGATGATGATG 1149
Db |||||
QY 1021 AACACATTTGATTCACCATTTGAAGCAGTCTGCACTTTCGCTTTCAATATGATGATGATG 1080
Db |||||
QY 1150 GTGACCTGCTCAAGATGCTTCCATTCCTGATGAGGATGCTGCTCCCACTGATCATT 1209
Db |||||
QY 1081 GTGACCTGCTCAAGATGCTTCCATTCCTGATGAGGATGCTGCTCCCACTGATCATT 1140
Db |||||
QY 1210 ACCCTCGGAGGCTGCTGTCGACACCGAGCTGCTGCAATGTTGATGATGATGATGATG 1269
Db |||||
QY 1141 ACCCTCAGAGGCTGCTGTCGACACCGAGCTGCTGCAATGTTGATGATGATGATGATG 1200
Db |||||
QY 1270 AAGTACATTTTGTGATCTGGGATGAACTATAAGGATGAGAAACAAAGTACTTGT 1329
Db |||||
QY 1201 AAGTACATTTTGTGATCTGGGATGAACTATAAGGATGAGAAACAAAGTACTTGT 1260
Db |||||
QY 1330 GAATTTGTAAGACCTTAATGACACAAACAGGAGCTGCTGTTGAGTACAGGAC 1389
Db |||||
QY 1261 GAATTTGTAAGACCTTAATGAGGACAGGCTGCTGTTGAGTACAGGAC 1320
Db |||||
QY 1390 AGGCTGTGTAGTGGCTCATCTGACAACTATCAGATTTATGGGACATAGATGTGT 1449

Db |||||
QY 1321 AGCTGTGTGAGCGCTCTCTGACAAACCATCAGGCTGTGGGACATAGATGTGGA 1380
Db |||||
QY 1450 GCATGTTTACGAGTGTAGAGGCCATGAGGAATTTGGTTCGTTTATTCGATTTGATAAC 1509
Db |||||
QY 1381 GCATGTTTACGAGTGTGAGGCCATGAGGAATTTGGTTCGTTTATTCGATTTGATAAC 1440
Db |||||
QY 1510 AAGGATGATGATGAGGCGCTATGATGGAATAATTAAGTGTGGATCTTATGCTGCT 1569
Db |||||
QY 1441 AAGGATGATGAGGCGCTATGATGGAATAATTAAGTGTGGATCTTATGCTGCT 1500
Db |||||
QY 1570 TTGAGCCCGCTGCTCTGAGGACACTCTGCTACGAGCCCTTTGTGAGCATTTCCGGA 1629
Db |||||
QY 1501 TTGAGCCCGCTGCTCTGAGGACACTCTGCTACGAGCACCTTTGTGAGCATTTCCGGA 1560
Db |||||
QY 1630 AGAGTGTTCGAGTACAGTTTCAATTCAGATTTGTCAGTGTTCACATGATGACACA 1689
Db |||||
QY 1561 AGAGTGTTCGAGTACAGTTTCAATTCAGATTTGTCAGTGTTCACATGATGACACA 1620
Db |||||
QY 1690 ATCTCATCTGGGACTTCTAAATGATCCAGCTGCCAAGCTGAACCCCGCTTCCCT 1749
Db |||||
QY 1621 ATTCTCATCTGGGACTTCTGAATGATCCAGCTGCTCAGCTGAACCCCGCTTCCCT 1680
Db |||||
QY 1750 TCTCGACATACACTTCTCCAGATAAT 1781
Db |||||
QY 1681 TCTCGACATACACTTCTCCAGATAAT 1712
Db |||||
RESULT 12
AF112979 1710 bp mRNA linear ROD 02-MAR-1999
LOCUS
DEFINITION Mus musculus beta-transducin repeat containing protein mRNA.
ACCESSION AF112979
VERSION AF112979.1 GI:4140717
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1710)
AUTHORS Spencer, E., Jiang, J. and Chen, Z. J.
TITLE Signal-induced ubiquitination of Ikappaalpha by the F-box protein
Slimb/beta-TrCP
JOURNAL Genes Dev. 13 (3), 284-294 (1999)
MEDLINE 99145465
PUBMED 9990853
REFERENCE 2 (bases 1 to 1710)
AUTHORS Chen, Z. J.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1998) Molecular Biology and Oncology, UT
Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX
75235-9148, USA
FEATURES
Location/Qualifiers
1..1710
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..1710
/function="ubiquitinates phosphorylated IkBa"
/notes="beta-TrCP; F-box protein; 1k8-ubiquitin ligase;
substrate recognition subunit of SCF complex; similar to
Homo sapiens beta-TrCP and Drosophila melanogaster Slimb"
/codon_start=1
/product="beta-transducin repeat containing protein"
/protein_id="AAD04181.1"
/db_xref="GI:4140718"

JOURNAL Mol. Cell. Biol. 13 (8), 4953-4966 (1993)
 MEDLINE 9330289
 PUBMED 8391141
 COMMENT Original source text: Xenopus laevis (library: S. cerevisiae
 expression library of X.laevis oocytes) cDNA to mRNA.
 FEATURES
 Location/Qualifiers
 1..1671
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /cell_type="oocyte"
 /tissue_lib="S. cerevisiae expression library of X.laevis
 oocytes"
 1..1671
 /gene="beta-TrCp"
 48..1604
 /gene="beta-TrCp"
 /codon_start=1
 /protein_id="AAA02810.1"
 /db_xref="GI:295543"
 /translation="MEGFSCSLQPPPTASERDCNRPPEPRKILTEKNTURQKLANGT
 SSMIVPORKLSANVEKEKELVKYFPMSECDQVEFVHLISRMCHYOHGHINYLK
 PMLORDETALPARGLDHIAENILSLYLDKSLCSAEILVCKEWVVTSDGLWKLLER
 MYRDSLRWGLAERGWQVLFKNKPPDGKTPNSFYRALYFKILODIETIESNWRCG
 RHLQRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIDWKMTLECKRVLMGHTGSLV
 CLOYDERVIITSSDSVTRVMDVNTGEMNLTHHCEAEVLHLRPNMMVTCSKDRSI
 AVMDASATDITLERVLVGHRAAVNVVDFDKIVISASGDRTIKVMNTTCEFTSLN
 GHKRGICLQVDRDLVVGSSDNTIRLWDIECGACLRVLEGHLELVRCTIRFONKFI
 GAYDGKIKVMDLVAALDPRAPAGTLCRLTLVHSGRVFLQDFEIVSSSHDDTILI
 WDFLNDPLGA"
 1671
 /gene="beta-TrCp"
 495 a 341 c 424 g 411 t
 polyA_site
 BASE COUNT 495 a 341 c 424 g 411 t
 ORIGIN
 Query Match 47.3%; Score 1017.8; DB 5; Length 1671;
 Best Local Similarity 81.8%; Pred. No. 2e-277;
 Matches 1175; Conservative 0; Mismatches 262; Indels 0; Gaps 0;
 284 AAACAAACTTGCCTTCCAGTATGATGTGCGCAAGCAAGCAAGCAACTCTGAG 343
 157 AGACAAACTTGCCTTCCAGTATGATGTGCGCAAGCAAGCAAGCAACTCTGAG 216
 344 CAAGCTATGAAAGGAAAGCAACTGTGTGTCCTTCAATATCTTGTGACGAGTGTGACAGTCTAG 403
 217 CAAATTACGAGAGGAAAGCAAGCAAGTATGCGTCAAGTATTTTGACGAGTGTGCGAGTGG 276
 404 ATCAAGTGAATTTGTGAAACATCTTATATCCCAATGTGTCTATACCAACATGAGGACACA 463
 277 ATCAAGTGAATTTGTGAAACATCTTATATCCCAATGTGTCTATACCAACATGAGGACACA 336
 464 TAAACTCGTATCTTAAACCTATGTTGACGAGAGATTTTCAATATGCTCTGCCAGCTCGGG 523
 337 TAAACACTTACCTTAAGCAATGTTACAAAGGATTTTCAATACCGCACTTACCACTCGG 396
 524 GATTGATCATATCGCTGAGAACATTTCTGTCACTACCTGGATGCGCAATCACTATGTGCTG 593
 397 GACTCGATCATACGAAATAATATCTTTCATACCTGGATGCGCAATCACTATGTGCTG 456
 584 CTGAACCTGTGTGCAAGGAATGTTACCGAGTGACCTCTGATGGCATCTGTGGAAGAAC 643
 457 CAGACTGGTATGTAAGAGTGTGTATCGAGTGACCTCAGATGGATGCTCTGGAAGAAC 516
 644 TTATCGAGAGATGTGAGGACAGATTTCTGTGAGAGGCTGTGCGCAAGCAAGAGGAT 703
 517 TCATAGAGCGGATGTGCGGACAGATTTCTTTGAGAGGAGCTGGCAGAAAGAGGAGAT 576
 704 GGGGACGATTTTATTTCAAAACAAACCTCTGACGGGATGTCTCTCCCAACTCTTTT 763
 577 GGGGCAATATTTGTTTAAACAAACCTCTGATGGAAACCCACCAATCTTATCT 636
 764 ATAGAGCACTTTATCTCTAAATTTATPACAGACATTTGAGACATAGATCTTAATTGGAGAT 823

637 ACAGAGCGCTTTACCCAAAAATTATTCAAGACATAGACAAATCGAGTCCAACTCGCGCT 696
 824 GTGGAAGACATAGTTTACAGAGAATTCACCTCCGAGGTGAACAAAGCAAGAGGATTTACT 883
 697 GTGGGAGACACAGCTTACAAAGAAATTCACCTCCGAGGTGAACAAAGCAAGAGGATTTACT 756
 884 GTTTACAGTATGATGATCAGAAATAGTAAAGGGGCTTCGAGACAAACACATCAAGATCT 943
 757 GTCTGAGTACGATGATCAGAAATAGTAAAGGGGCTTCGAGACAAACACATCAAGATCT 816
 944 GGGATAAAAACACATTTGGAATCAAGCGAATTTCTACAGGCGCATACAGGTTTCAGTCTCT 1003
 817 GGGATAAAGAAATCTTTGGAGTCAAGCGAGTGTGATGGGTCAACATGGGTCACTGGGTCTCT 876
 1004 GTCTCCAGTATGATGAGAGAGTATCAATCAAGGATCATCGGATTCACCGTCCAGAGTGT 1063
 877 GTCTGCAATATGATGAGAGAGTATCAATCTAGTGGCTCTTCAGATCTCTACCGTCCGGGTGT 936
 1064 GGGATTAATACAGGTGAATGCTTAAACACAGTGTGATTCACCATTTGTGAAGCAGTTCTGTC 1123
 937 GGCACGTGAACACAGGAGAAATGTTGAACACGCTGATTCACCATCTGTGAGGCTGTGCTGC 996
 1124 ACTTGCGTTTCAATATGCGATGATGCTGCTGCTCAAGAGATCGTTCATTCCTGCTGAT 1183
 997 ACTTGAGGTTTAAATATGCGATGATGCTGCTGCTCAAGAGATCGTTCATTCCTGCTGAT 1056
 1184 GGGATATGCGTCCCCCACTGACATTTACCTCCGAGGAGTGTGCTGCTGCGACACCGAGCTG 1243
 1057 GGCACATGCGCTCTGCGACCGATATCACATTACGAAGAGTTTCTGGTAGGCCACCGAGCTG 1116
 1244 CTGCTCAATGTTGAGACATTTGATGACAGTACATTTGTTCTGCTGCTGCGGATGAGACTA 1303
 1117 CCGTAAATGTGGTGGACTTTGATGACAAATATAGTCTCTGCTGCTGCTGATGATCAACAA 1176
 1304 TAAAGGTATGGAACACACAGTACTTGTGAATTTGTAAGGACCTTAAATGAGACACAAACGAG 1363
 1177 TAAAGGTTTGGAAACACCACTATCATGTGAATTTGTGCGGACATTTGAACGCGCCACAGCGT 1236
 1364 GCATTCGCTTGTGCGTACAGGACAGCTGTTGATGAGTGGTGGCTCATCTGACAACTA 1423
 1237 GTATTGCTGCTTACAGTATCGAGATCGGCTGTTGCTGAGTGGTCTTCTGACAACTA 1296
 1424 TCAGATTATGGGACATAGAAATGTGTGATGTTTACGAGTGTGTAAGAGCCATCAGGAAT 1483
 1297 TCAGATTGCGGATTTGAATCGGTGATGTTTTCGCGGTTCTGGAAGGACATGAAGAAC 1356
 1484 TGGTGGTGTGATTTGATTTGATTAACAGAGATGATCAGTGGGCTTATGATGAGAAA 1543
 1357 TGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416
 1544 TTAAGTGTGGGATCTTGTGGTGTGCTTTGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603
 1417 TTAAGTGTGGGATCTTGTGGTGTGCTTTGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1476
 1604 TAGGACCTTGTGGAGCATTCGGAAGAGTTCGAGTACAGTACAGTACAGTACAGTACAGTACAG 1663
 1477 TCGGAGCTTGTGGAGCATTCGGAAGAGTTCGAGTACAGTACAGTACAGTACAGTACAGTACAG 1536
 1664 TTGCTAGTATGATCAGATGACAAATCTCTCATCTGGGACTTCTTAATGATCCAG 1720
 1537 TAGTAAGCAGCTTCCAGATGACATATCTCTCATCTGGGATTTTCTCAACGACCCAG 1593
 RESULT 14
 AB033279
 LOCUS
 DEFINITION Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
 beta-TRCP2 isoform A, complete cds.
 ACCESSION AB033279
 VERSION AB033279.1
 KEYWORDS BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform A.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)
Koike.J., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M. and
Katoh,M.
Molecular cloning and genomic structure of the betaTRCP2 gene on
chromosome 5q35.1
Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)

20160458
10694485
2 (bases 1 to 2134)
Katoh,M.
Direct Submission
Submitted (05-OCT-1999) Masaru Katoh, National Cancer Center,
Genetics Division; Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
(E-mail: mkatoh@ncc.go.jp, Tel:81-3-3542-2511 (ex.4402),
Fax:81-3-3541-2685)

Location/Qualifiers
1. .2134
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q35.1"
/tissue_type="fetal lung"

1. .2134
/gene="BTRCP2"
102..1628
/gene="BTRCP2"
/codon_start=1
/product="F-box and WD-repeats protein beta-TRCP2 isoform
A"

/protein_id="BAA92329.1"
/db_xref="GI:7209809"
/translation="MEPDSVIEDKTIELMSNGTSVIVSRKPSSEGVYQKDLCLIK
YFDWSSDQVEFVHLKISMCHYQHGHSYKLPMLQDFITALEPQGLDRIENIL
SYLDARSLCAELVCKEWQVISEGLMKLIERMVRTDPLMKVSGSERGWDQYLKFN
RPTDGPNSFYRSLPKIIDIETIESNMRGRHNLQRIQCRSENSKGVYCIQYDDBK
IISLRNSFKIMDKTSLKVLGTHTGSLVLCLOYDERIVITGSSDSVTVRVMDVNTG
EVNLTLLHNEAVLHFLFNGSLMTGSKRSIAVDMASATDITLRRVLVGHRAAVV
VDPDKIVASAGDRTIKWSTSTCEPVRTLNHKGRIACLOVRDLRVVSGSSDNTIR
LWDTECCACLVRLGHELVRCIRFDPKRIVSGAYDGKIKWDLQALDPRAPASTLC
LRLTVEHSGRVRFLQFDEPQLISSHDDTLIWDVFLNVPFSAQNETRSPRTYTIYSR
"

BASE COUNT 566 a 477 c 515 g 576 t
ORIGIN

Query Match 41.1%; Score 883.4; DB 9; Length 2134;
Best Local Similarity 74.8%; Pred. No. 3.2e-239;
Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

2y 298 AATGGCACTTCATATGCCAATGTGTCATCAACATGGGCACATFAAACTCGTATCTT 477
|||||
db 153 AATGGAACATCATCTGTGATCGTCTCCAGAAAGAGGCCATCAGAAAGAACTATCAAAA 212
|||||

2y 358 GAAAGGAACGTGTGTCARATCTTTGAGCAGNGGTTCAGATCAGATCAAGTGGATTT 417
|||||
db 213 GAAAGAACTGTGTATTAATAATTTTGACCAAGTGGTCTGAATCAGATCAAGTGGATTT 272
|||||

2y 418 GTGGAACATCTTATATCCAAATGTGTCATTAACCAATGGGCACATFAAACTCGTATCTT 477
|||||
db 273 GTGGAACATCTTATTCAGGAATGTGTCATATCAGATCGACATATTAATCTTACCTG 332
|||||

2y 478 AAACCTATGTTGCAGAGAGATTCATTAACCTCTCGCAGCTGGGATGGATCATATC 537
|||||
db 333 AAGCCCATGTTGCAGCGGACTTTTATACCGCTTTACAGAGCAAGGCTTAGATCATATA 392
|||||

2y 538 GCTGAGAACATCTGTCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACCTTGTGTC 597
|||||
db 393 GCAGAAACATCTCTTCTACTCTGGATGCCAGGTCTCTGTGTCAGACAGACTGGTATGT 452
|||||

2y 598 AAGGAATGGTACCGAGTACCTCTGATGGCATGCTGTGGGAAGAGCTTATCGAGAGATG 657
|||||
db 1527 CATGATGACACTATTTTGGATTTGGGATTTCTTAAATGTGCTCCAGTCCCAAGATGAG 1586
|||||

453 AAGAAATGGCAGCGAGTGATCTCAGAAAGNATGCTTTGGAAGAGCTGATTGAACGAATG 512
|||||

Qy 658 GTCAGGACAGATTCTCTGTGGAGAGCGCTGGCAGAACGAGAGGATGGGACAGTATTTA 717
|||||

Db 513 GTACGCACCTGATCCCTTATGGAAGGACTTTCAAGAAAGAGAGGTGGGATCAGTACCTG 572
|||||

Qy 718 TTCAAAACAAACCTCTCTGACGGGATGCTCTCCCAACTCTTTTATATAGACACTTTAT 777
|||||

Db 573 TTTAAAACAGACCCACAGATGGC-----CTTCCAAATTCATTTTATAGGTCATTATAC 626
|||||

Qy 778 CCTAAAATATACAAGACATTGAGACAATAGATATCTAATTTGGAGATGTGAAGACATAGT 837
|||||

Db 627 CCAAGATTATTCAGGATATAGAGACTATAGATCTAACTGGCGGTGTGGACAGACAAC 686
|||||

Qy 838 TTACAGAGATTTACTGCCAGAGTGAACAGCAAGAGGTTTACTGTTTACAGTATGAT 897
|||||

Db 687 TTGAGAGGATTCAGTGCCTCTGAAATAGTAAAGGTGTCTACTGTTTACAGTACGAT 746
|||||

Qy 898 GATCAGAAAATAGTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATAAAAACACA 957
|||||

Db 747 GATGAAAAATTTATCAGTGGCTACGAGATAATCTTATTAAGATATGGGATAAACCAGC 806
|||||

Qy 958 TTGAATGCAAGCGAATCTCAAGGCCATACAGTTTCAGTCTCTGCTCCAGTATGAT 1017
|||||

Db 807 CTGGAATGTTTGAAGGTGTTAAACAGGACACACAGGCTCTGCTCTGTGTCAGTATGAT 866
|||||

Qy 1018 GAGAGAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGTGGGATGTAATACA 1077
|||||

Db 867 GAGCGTCTCATTTGAACTGGCTCTTCAGATTCTACCGTGAGAGTGTGGGATGTGAACAG 926
|||||

Qy 1078 GGTGAATGTAAACAGTTGATTCACATTTGGAAGCAGTTCTTGCACTTGGCTTTCAAT 1137
|||||

Db 927 GGTGAAGTCTTAAACATTTGATCCACCAATAGAGCTGTATTGCACTTAGCTTCAGC 986
|||||

Qy 1138 AATGGCATGTGTGACCTGCTCAAGATCGTTCATTCCTGTATGGGATATGGCTCC 1197
|||||

Db 987 AATGGCATGTGTGACCTGCTCAAGAGCCGCTCCATTCCTGTGTGGGACATGGCTTCT 1046
|||||

Qy 1198 CCAACTGACATTTACCTCCCGAGGGTGTGTGGGACACCGAGCTGCTGTCAATGTTGTA 1257
|||||

Db 1047 GCGACCCACATCACTTTACCGCTGTCTGTGGTGGCCACCGGCTCCCGCTCAATGTAGTA 1106
|||||

Qy 1258 GACTTTGATGACAAAGTACATGTTTCTGCACTGGGGATAGAACTATAAAGGTATGGAAAC 1317
|||||

Db 1107 GACTTTGACGACAAAGTACATGTTGCTGCTGCTGGTGTGACAGGACCATCAAGCTCTGGAGC 1166
|||||

Qy 1318 ACAAGTACTTGTGAATTTGTAAGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTG 1377
|||||

Db 1167 ACAGACACTGTGAAATTTGTTGCTACTCTCAATGGGACACAGGGGGCATTCGCTGTCTC 1226
|||||

Qy 1378 CAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGACAAACATATCAGATATGGGAC 1437
|||||

Db 1227 CAGTACAGGATCGCTGTTGTTAGTGGATCATCAGATAATACCAATAGGCTCTGGAT 1286
|||||

Qy 1438 ATGAAATGTGTGTCATGTTTACGAGTGTGTAAGGCCATGAGGAATTTGTGCTGCTTATT 1497
|||||

Db 1287 ATTGAATGTGTGCTGCTTTTAAAGTCTTAGAGGACATGAAGATTTGTCGATGCATC 1346
|||||

Qy 1498 CGATTTGATTAACAGAGGATAGTCACTGGGGCCCTATGATGGAATAAATTAAGTGTGGGAT 1557
|||||

Db 1347 CGGTTTGTAAACAGAGGATTTGTCAGTGGGGCCCTATGATGGGAAATTAAGATTTGGGAC 1406
|||||

Qy 1558 CTTGTGCTGCTTTGGACCCCGTCTCTGAGGAGACACTCTGCTACGGACCCCTTGTG 1617
|||||

Db 1407 TTGAAGCTGCTCTTGACCTCGAGCCCGCAGCAGACACATTTGTGTTTGGCCATTTGGTG 1466
|||||

Qy 1618 GAGCATTTCCGGAAGAGTGTTCGACTACAGTTTGAATGATGAAATTCAGATTTGTCAGTATGCA 1677
|||||

Db 1467 GAAATTTCTGACGTGTGTTTGGCTCCAGTTTGTAGTGTTCAGATCATCAGACACTCC 1526
|||||

Qy 1678 CATGATGACACAATCTCTCATCTCGGACTTCTTAATGATCCAGCTGCCAAGCTGAACCC 1737
|||||
Db 1527 CATGATGACACTATTTTGGATTTGGGATTTCTTAAATGTGCTCCAGTCCCAAGATGAG 1586
|||||


```

2Y 1738 CCCGTTCCCTCTCGAACATACACCTACATCTCCAGATTAATACCATACACTGACC 1796
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1587 ACCCGTTCCCTCCAGAACATACACTTACATCTCTAGATAACAGTCTGCACCTTCCACC 1645

RESULT 15
LOCUS AB033280 2252 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
ACCESSION AB033280
VERSION beta-TRCP2 isoform B, complete cds.
KEYWORDS BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform B.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Koike,J., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M. and
Katoh,M.
Molecular cloning and genomic structure of the betaTRCP2 gene on
chromosome 5q35.1
JOURNAL Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
MEDLINE 20160458
PubMed 10594485
REFERENCE 2 (bases 1 to 2252)
AUTHORS Katoh,M.
DIRECT SUBMISSION
TITLE Submitted (05-OCT-1999) Masaru Katoh, National Cancer Center,
Genetics Division, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
(E-mail:mkato@ncc.go.jp, Tel:81-3-3542-2511(ex.4402),
Fax:81-3-3541-2685)
FEATURES
    Location/Qualifiers
        1..2252
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="5"
            /map="5q35.1"
            /tissue_type="fetal lung"
        1..2252
            /gene="BTRCP2"
            /locus_tag="BTRCP2"
            /gene="BTRCP2"
            /codon_start=1
            /product="F-box and WD-repeats protein beta-TRCP2 isoform
            B"
            /protein_id="BAA92330.1"
            /db_xref="GI:7209811"
            /translation="MEPDSVIEDKTIELMNTSVMEQNEDESPKKNLWQISNGTSSV
            IVSRKPSGNYQKEDLCIKYFDQSESDQVEFVHLISRMCHVQHGHINSYLPML
            ORDFTLALPEOGLDHIENILSYLDARSLCAELVCKEWRVISEGMLWKLIERWVR
            TDLWKLGLSERGWDVLFKNRPDGPNSFVSLPKLIQDLETIESNWRGRENLO
            RLOCRSNSKGVCLQYDDEKILSGLDNSIKIWDKTSLECLVLTGHGVSILQYD
            ERVIVTSSDSTVAVWNTGEVNTLIHNEAVLHLRFSNGLMVTCSKDRSIAWDM
            ASADITLRLVLVGHRAVNVDPFDKIYIVSASGDRTIKVWSTCECFVFTILNGHRG
            IACIQYDRDLVSSSSNTIRLNDIBCCACLRVLEGHELVIRCFDNRKRIYVSGAYDG
            KIKWMDLQALDDPRAPASTLCRLTLVHSGRVRFLQDFBQIISSHDDTILWDFLN
            VPPGAQNETRSPSTYTYISR"
        598 a 501 c 558 g 595 t
BASE COUNT
ORIGIN
Query Match 41.1%; Score 883.4; DB 9; Length 2252;
Best Local Similarity 74.8%; Pred. No. 3.2e-239;
Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

2Y 298 AATGGCACTTCAGTATGTCGCCAAGCAACGGAACCTCTCAGCAAGCTTGAAGAG 357
Db 271 AATGAAACATCATCTGTGATCGTCTCCAGAGAGGCGCATCAGAGGAACATCAAAA 330

358 GAAAGGAACGTGTGTCAAACTCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGAATTT 417
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 331 GAAAAAGACTGTGTATTAATATTTTGGACAGTGGTCTGTAATCAGATCAAGTGGAAATTT 390
QY 418 GTGGAACATCTTATATFCCAAATGTGTCAATACCAACATGGGCACATAAACTCGTATCTT 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 GTGGAACATCTTATTTTCAGAAATGTGTCAATATCAGCATGGACATATTAATCTTTACCTG 450
QY 478 AAACCTATGTTCAGAGAGATTTTCACTACTGTCGCCAGCTCGGGATTCGGATCATATC 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AAGCCCATGTTGAGGGGACCTTTATACCGCTTTACAGAGCAAGGCTTAGATCAATATA 510
QY 538 GCTGAGAACATTTCTGTATCATCTGGATGCCAAATCACTATGTGTCTGTGAACCTTTGTGTC 597
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 GCAGAAAAACATTTCTTTCGTACCTGGATGCCAGGTCTCTGTGTGCAGCAGAGCTGGTATGT 570
QY 598 AAGGAATGTGTACCGAGTGACCTCTCATGTGSCATGCTGTGAGAGAGCTTTATCGAGAGATG 657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 AAGAATGCGAGGAGTGATCTCAGAAGAAATGCTTTGGAAGAGCTGATTTGAACGAATG 630
QY 658 GTCAGACAGATTTCTCTGTGGAGAGCGCTTGCAGAACCAAGAGGATGGGACAGTATTTA 717
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 GTACGCACCTGATCCCCATATGGAAGAGCACTTTTCAGAAAGAGAGGCTGGATCAGTACCTG 690
QY 718 TTCAAAAAACAAACCTCTCTGACGGGAATGCTCTCCAACTCTTTTATAGAGCACTTTAT 777
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 TTTAAAAACAGCCACAGATGCGC-----CCTCCAAATTCATTTATAGCTCATTTATAC 744
QY 778 CTTAAATTTATACAGACATTTGAGACAATAGATATCTAATTTGAGATGTGGAGACATAGT 837
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 745 CCAAGATTTATCAGGATATAGAGACTATAGAACTTAATCTGCGCGTGTGGACGACACAAC 804
QY 838 TTACAGAGAAATTCACCTGCGGAGTGAACCAAGCAAGGAGTTTACTGTTTACAGTATGAT 897
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 805 TTTCAGAGGATTCAGTGCCTCTCTGAAAATAGTAAAGTGCTCTGTTTACAGTACGAT 864
QY 898 GATCAGAAAATAGTAGCGGCTTCGAGACACACAANTCAAGATCTGGATTAATAACACA 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 865 GATGAAAAAATTTATCAGTGGCTTACGAGATAATTTCTATTAAGATATGGATTAATAACACG 924
QY 958 TTGGAATGCAAGGAAATCTCAGAGCCATACAGGTTTCAGTCTCTGTCCTCCAGTATGAT 1017
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 925 CTGGAATGTTTGAAGTGTAAACAGGACACACAGGCTCTGTCTCTGTCGCAATGAT 984
QY 1018 GAGAGAGTGCATCATACAGGATCATCGGATTCACCGTTCACAGTGTGGATGTAAATACA 1077
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 985 GAGCGTGTCAATGTAACTGGCTCTTCAGATTCTACGGTGAGAGTGTGGATGTGAACACG 1044
QY 1078 GGTGAATGCTAAACAGCTTGATTCACCATTTGTGAGCAGTCTTCTGCACTTCGGTTCAAT 1137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1045 GGTGAAGTTCCTTAACACATTTGATCCACCAATAGGCTGTATTGCACTTACGCTTCAGC 1104
QY 1138 AATGGCATGATGGTACCTGCTCCAAAGATCGTTTCATTGCTGTATGGATATGGCTCC 1197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1105 AATGGACTGATGTGACCTGTTTCCAAAGACCGCTCCATTGCTGTGTGGACATGGCTTCT 1164
QY 1198 CCAACTGCATTAACCTCCGAGGGTGTGGTTCGACACCGAGCTGCTGTCAAATGTTGTA 1257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1165 GCAGCTGCATCACTTTACGCCGTCTCTGTTGTGCGCCCGGGCTGCCGTCAATGTAGTA 1284
QY 1258 GACTTTTGATGACAAGTACATTTGTTCTGCATCTGGGGATAGAACTATATAAGGTATGGAAC 1317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1225 GACTTTGACGACAAGTACATCGTGTCTGCCCTGTGTGACAGACCATCAAAAGCTGGAGC 1284
QY 1318 ACAAGTACTTGTGAATTTGTAAAGGACCTTAAATGGACACAACGAGGATTCCTGTTTG 1377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1285 ACAGACACCTGTGAATTTGTTCTGTTACTCTCAATGGGCACAAGCGGGGCAATTCCTGTC 1344
QY 1378 CAGTACAGGACAGGCTGGTGTAGTGGCTCATCTGCAACACATCATCAGATTTATGGAC 1437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1345 CAGTACAGGATCGCTGTTGTTTGTAGTGATCATCAGATATATACATTTAGGCTCTGGAT 1404
QY 1438 ATAGAATGTTGTCATGTTTACAGTGTTAGAGGCCATGAGGAATTCGTCGTTGATTT 1497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1405 ATTGAATGTGGTCTGTTTAAAGATCTCTAGAGGACATGAAGAATTTGGTCCGATGCATC 1464

```



```

y 1498 CGATTGATACAAAGAGAGATAGTCAGTGGGGCCCTATGATGGAAAAAATTAAAGTGGGAT 1557
b 1465 CGGTTTGATACAAAGAGAGATTGTCACTGGGGCCCTATGATGGAAAAAATTAAAGTGGGAC 1524
y 1558 CTTGTGGCTGCTTTGGACCCCGTCTCTGCGAGGGACACTCTGTCTACGGACCCCTTGTG 1617
b 1525 TTGCAAGCTGCTCTTGACCTCGAGCCCGAGAGCAGACATGTGTGTCGACATGGTG 1584
y 1618 GAGCATTCCGGAAGAGTTTTTCGACTACAGTTTGGATGAAATTCAGATTGTCAGTAGTTCA 1677
b 1585 GAACATTCTGACGCTGTGTTTCGGCTCCAGTTTGATGAGTTTCAGATCATCAGCAGCTCC 1644
y 1678 CATGATGACACAACTCTCATCTGGGACTTCCTAAATGATCCAGCTGCCCAAGCTGACCC 1737
b 1645 CATGATGACACTATTTTGATTGGGATTTCTTAAATGTGCTCCAGTCCCAGAGATGAG 1704
y 1738 CCCGTTCCCTCTCGAACAATACACCTACATCTCCAGATAAATAACCATACACTGACC 1796
b 1705 ACCCGTTCTCCCTCCAGAACATACACTTACATCTCTAGATAACAGTCTGCACCTTTCACC 1763

```

search completed: October 22, 2003, 20:48:52
ob time : 7717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: October 22, 2003, 18:18:24 ; Search time 4097 Seconds
(without alignments)
12760.295 Million cell updates/sec

Title: US-09-601-168B-1
Perfect score: 2151
Sequence: 1 tgccttggtcgcgcctggc.....gtttgccagaaaaaaa 2151

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_pig.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1591.2	74.0	2897	11 AK032221	AK032221 Mus muscu
2	1557.2	72.4	2970	11 AK083257	AK083257 Mus muscu
3	1263.8	58.8	4502	11 AK052317	AK052317 Mus muscu
4	877.8	40.8	3824	11 AK041532	AK041532 Mus muscu

5	638.6	32.5	988	13	EX398796	EX398796
6	650.6	30.2	709	10	BE782628	BE782628
7	643.4	29.9	776	14	CR244526	CR244526
8	638	29.7	778	13	BQ444188	BQ444188
9	635.2	29.5	652	9	AL040095	AL040095
10	625	29.1	851	13	BQ425358	BQ425358
11	624.8	29.0	973	12	BM465280	BM465280
12	622.2	28.9	887	13	BUS07728	BUS07728
13	616.6	28.7	744	14	CB520936	CB520936
14	613.2	28.5	758	10	BE296484	BE296484
15	608.8	28.3	692	10	BE269297	BE269297
16	603.2	28.0	749	14	CR246650	CR246650
17	589.4	27.4	828	12	B1822845	B1822845
18	586.6	27.3	1147	13	BQ948186	BQ948186
19	580.6	27.0	743	12	BG829037	BG829037
20	579.4	26.9	895	9	AL898993	AL898993
21	574.2	26.7	699	14	CR246901	CR246901
22	566.4	26.3	752	10	BG722472	BG722472
23	561.6	26.1	1136	14	CD500760	CD500760
24	557	25.9	779	14	CD350839	CD350839
25	553.6	25.7	924	14	CA985560	CA985560
26	540	25.1	870	13	BQ463637	BQ463637
27	534.4	24.8	789	10	BG518761	BG518761
28	534.2	24.8	966	14	CA980902	CA980902
29	532.6	24.8	555	14	CA416190	CA416190
30	532.4	24.8	637	12	B1103092	B1103092
31	529.8	24.6	544	9	AL041438	AL041438
32	528.4	24.6	888	12	B1455545	B1455545
33	521.4	24.2	729	14	CB948025	CB948025
34	518	24.1	756	13	BQ460380	BQ460380
35	516.2	24.0	810	12	BM944304	BM944304
36	513.6	23.9	846	9	AL879775	AL879775
37	511.6	23.8	636	13	BU055333	BU055333
38	500.4	23.3	646	10	BG082149	BG082149
39	497	23.1	850	13	BQ771102	BQ771102
40	494.8	23.0	850	14	CA327271	CA327271
41	494.6	23.0	1002	14	CA971789	CA971789
42	492.2	22.9	541	10	BG058605	BG058605
43	488.8	22.7	590	9	AW762522	AW762522
44	488.2	22.7	806	13	BU750387	BU750387
45	488	22.7	488	13	EX099276	EX099276

ALIGNMENTS

RESULT 1	AK032221	2897 bp	linear	HTC 05-DEC-2002
LOCUS	AK032221	Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430504E23 product:beta-transducin repeat containing protein, full insert sequence.		
DEFINITION	AK032221			
ACCESSION	AK032221			
VERSION	AK032221.1	GI:26082733		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1			
AUTHORS	Carninci,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,N., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2970)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Riken Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1..2970
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C630031K01"
/db_xref="taxon:10090"
/clone="C630031K01"
/sex="male"
/tissue_type="hippocampus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
81..1900
/note="beta-transducin repeat containing protein (MGDI|MGDI:1338871, GB|NM_009771, evidence: BLASTN, 100%, match=1502)
putative"

BASE COUNT 739 a 758 c 788 g 685 t

ORIGIN

Query Match 72.4%; Score 1557.2; DB 11; Length 2970;
Best Local Similarity 88.2%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 218; Indels 22; Gaps 9;

112 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 171
113 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 230
114 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 290
115 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 350
116 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 410
117 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 470
118 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 530
119 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 590
120 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 650
121 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 710
122 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 770
123 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 830
124 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 890
125 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 950
126 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1010
127 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1070
128 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1130
129 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1190
130 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1250
131 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1310
132 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1370
133 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1430
134 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1490
135 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1550
136 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1610
137 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1670
138 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1730
139 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1790
140 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1850
141 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1910
142 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 1970
143 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2030
144 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2090
145 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2150
146 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2210
147 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2270
148 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2330
149 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2390
150 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2450
151 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2510
152 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2570
153 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2630
154 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2690
155 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2750
156 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2810
157 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2870
158 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2930
159 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 2990
160 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3050
161 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3110
162 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3170
163 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3230
164 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3290
165 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3350
166 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3410
167 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3470
168 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3530
169 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3590
170 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3650
171 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3710
172 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3770
173 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3830
174 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3890
175 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 3950
176 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4010
177 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4070
178 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4130
179 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4190
180 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4250
181 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4310
182 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4370
183 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4430
184 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4490
185 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4550
186 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4610
187 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4670
188 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4730
189 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4790
190 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4850
191 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4910
192 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 4970
193 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5030
194 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5090
195 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5150
196 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5210
197 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5270
198 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5330
199 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5390
200 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5450
201 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5510
202 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5570
203 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5630
204 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5690
205 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5750
206 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5810
207 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5870
208 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5930
209 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 5990
210 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6050
211 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6110
212 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6170
213 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6230
214 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6290
215 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6350
216 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6410
217 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6470
218 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6530
219 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6590
220 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6650
221 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6710
222 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6770
223 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6830
224 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6890
225 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 6950
226 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7010
227 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7070
228 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7130
229 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7190
230 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7250
231 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7310
232 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7370
233 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7430
234 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7490
235 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7550
236 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7610
237 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7670
238 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7730
239 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7790
240 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7850
241 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7910
242 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 7970
243 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8030
244 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8090
245 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8150
246 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8210
247 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8270
248 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8330
249 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8390
250 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8450
251 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8510
252 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8570
253 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8630
254 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8690
255 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8750
256 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8810
257 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8870
258 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8930
259 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 8990
260 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9050
261 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9110
262 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9170
263 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9230
264 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9290
265 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9350
266 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9410
267 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9470
268 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9530
269 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9590
270 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9650
271 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9710
272 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9770
273 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9830
274 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9890
275 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 9950
276 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10010
277 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10070
278 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10130
279 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10190
280 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10250
281 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10310
282 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10370
283 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10430
284 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10490
285 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10550
286 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10610
287 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10670
288 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10730
289 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10790
290 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10850
291 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10910
292 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 10970
293 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11030
294 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11090
295 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11150
296 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11210
297 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11270
298 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11330
299 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11390
300 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11450
301 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11510
302 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11570
303 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11630
304 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11690
305 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11750
306 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11810
307 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11870
308 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11930
309 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 11990
310 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12050
311 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12110
312 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12170
313 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12230
314 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12290
315 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12350
316 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12410
317 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12470
318 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12530
319 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12590
320 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12650
321 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12710
322 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12770
323 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12830
324 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12890
325 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 12950
326 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13010
327 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13070
328 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13130
329 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13190
330 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13250
331 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13310
332 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13370
333 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13430
334 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13490
335 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13550
336 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13610
337 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13670
338 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13730
339 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13790
340 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13850
341 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13910
342 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 13970
343 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14030
344 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14090
345 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14150
346 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14210
347 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14270
348 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14330
349 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14390
350 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14450
351 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14510
352 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14570
353 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14630
354 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14690
355 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14750
356 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14810
357 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14870
358 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14930
359 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 14990
360 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15050
361 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15110
362 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15170
363 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15230
364 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15290
365 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15350
366 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15410
367 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15470
368 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15530
369 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15590
370 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15650
371 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15710
372 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15770
373 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15830
374 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15890
375 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 15950
376 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16010
377 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16070
378 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16130
379 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16190
380 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16250
381 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16310
382 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16370
383 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16430
384 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16490
385 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16550
386 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16610
387 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16670
388 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16730
389 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16790
390 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16850
391 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16910
392 TTTATGAAATCTTCAGAGAGAGACACTGTATATATGCGGACCCCTTAGAGATTAATA 16970

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source 1. .4502
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D330027D11"
/db_xref="taxon:10090"
/clone="D330027D11"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
misc_feature 1. .4502
/note="beta-transducin repeat containing protein (MGDI|GI:1338871, GB|NM_009771, evidence: BLASTN, 100%, match=1502)"

BASE COUNT 1284 a 998 c 1126 g 1094 t
ORIGIN

Query Match 58.8%; Score 1263.8; DB 11; Length 4502;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

```

2y 112 TTTATCAATTCTCTCAGAGAGAGACCTGTAATATGCGAACCCCTAGGAGATAATA 171
Db 280 TTCCAGAAATCTTCAGAGAGAGACCTGTAATATGCGAACCCCTAGGAGATAATA 339
2y 172 CCAGAGAGAAATTCATTAGACAGACATACACAGCTGTGCAGACTCTGTTAAACCAA 231
Db 340 CCAGAGAGAAATTCATTAGACAGACTTACACAGCTGTGCAGCTTTGCATAAACCAA 399
2y 232 GAAACAGTAGTTTACAGACACTGCTATGAAGACTGAGACTGTTGTGGCCMAACAAA 291
Db 400 GAGACAGTAGTTTACAGACACTGCTATGAAGACTGAGAAATTTGTGTGGCCMAACCAA 459
2y 292 CTTGCCAATGGCACTTCCAGTAGTATGTTGCCAAGCAACGGAACCTCTCAGCAAGCTAT 351
Db 460 CTTGCCAATGGCACTTCCAGTAGTATGTTGCCAAGCAACGGAACCTCTCAGCAAGCTAT 519
2y 352 GAAAGGAAAGAACTGTGTCAAACTTTTGAGCAGTGTGAGAGTCAAGTCAAGTG 411
Db 520 GAGAGGAAAGAAAGTGTGTCAAGTATTTTGAGCAGTGTGAGAGTCAAGTCAAGTG 579
2y 412 GAATTTGTGGAACATCTTATATCCCAATGTGTCTATACCAACATGGGCACATAACTCG 471
Db 580 GAATTTGTGGAACATCTTATATCCCAATGTGTCTATACCAACATGGGCACATAACTCC 639
2y 472 TATCTTAAACCTATGTCAGAGAGATTCATAACTGCTCGCCAGCTCGGGGATTGGAT 531
Db 640 TACTTAAACCTATGTCAGAGAGATTCATAACTGCTCGCCAGCTCGGGGATTGGAT 699
2y 532 CATATCGCTGAGAACATCTGTCTACTCTGGAATGCGAAATCACTATGCTGTGAACTT 591
Db 700 CATATCGCTGAGAACATCTGTCTACTCTGGAATGCGAAATCACTATGCTGTGAACTT 759
2y 592 GTGTGCAAGGATGTGTACCGAGTGTGCTGTGATGCTGTGGAAGACTTATCGAG 651
Db 760 GTGTGCAAGGATGTGTACCGAGTGTGCTGTGATGCTGTGGAAGACTTATCGAG 819
2y 652 AGAATGTCAGGACAGATCTCTGTGGAAGAGCCCTGGCAGAACGAGAGGATGGGACAG 711
Db 820 AGAATGTCAGGACAGATCTCTGTGGAAGAGCCCTGGCAGAACGAGAGGATGGGACAG 879
2y 712 TATTTATTCAAAACAAACCTCTCTGACGGGATGCTCTCCCAACTCTTTTATAGAGCA 771
Db 880 TACTTATTCAAAACAAACCTCTCTGATGAGAACGCTCTCTCCCAACTCTTTTATAGAGCG 939

```

```

Qy 772 CTTTATCTTAAAAATTATACAAGACATTTGAGACAATAGAAATCTAATTGGAGATGTGAAGA 831
Db 940 CTTTATCTTAAAAATCATACAAGACATTTGAGACAATAGAAATCTAATTGGAGATGTGAAGA 999
Qy 832 CATAGTTTACAGAGAAATTCACCTGCCGAGTGAACCAAGCAAGAGGAGTTTACTGTTTACAG 891
Db 1000 CATAGTTTACAGAGAAATTCACCTGCCGAGTGAACCAAGCAAGAGGAGTTTACTGTTTACAG 1059
Qy 892 TATGATGATCAGAAAAATAGTAAGACGCTTTCAGAGCAACACAAATCAAGATCTGGGATAAA 951
Db 1060 TACGACGACGAGAGATAGTCTCAGCGGCTTTCAGAGCAACACCAATCAAGATCTGGGATAAA 1119
Qy 952 AACACATTTGGAAATGCAAGCGAATTTCTCAGAGCCCATACAGGTTTCACTCTCTCTCAG 1011
Db 1120 AGCACACTGGAAATGCAAGCGGATTTCTCAGCGGCCACACGGGCTCCGCTCTGTCTCTGCG 1179
Qy 1012 TATGATGAGAGATGATCATATAACAGGATCATCGGATTCACGGTCCAGAGTGTGGGATGTA 1071
Db 1180 TACGATCAGAGGCTGATCATCATCAGGCTCTCAGACTCCACCGTCAGAGTGTGGGATGTA 1239
Qy 1072 AATACAGGTCGAATGCTAAACACGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTTGGCT 1131
Db 1240 AATCAGGTCGATGCTAAACACATTTGATTCACCATTTGTGAAGCAGTTCTGCACCTTGGCT 1299
Qy 1132 TTCAATATGGCATGATGTGACCTCTCCAAAGATCGTTCCATTTGCTGTATGGGATATG 1191
Db 1300 TTCAATATGGCATGATGTGACCTCTCCAAAGACCGTTCCATCGCTGTGTGGGATATG 1359
Qy 1192 GCCTCCCACTACATTTACCTCCGAGGCTGTGTCGACACCGAGCTGTGCTCAAT 1251
Db 1360 GCTTCCCACTACATTTACCTCCGAGGCTGTGTCGACACCGAGCTGTGCTCAAT 1419
Qy 1252 GTTGATAGCTTTGATGACAAAGTACATTTCTGCAATCTGGGGATAGAACTATAAAGTGA 1311
Db 1420 GTTGATAGCTTTGATGACAAAGTACATTTCTGCTCTCGGAGATAGAACCAATAAGGTG 1479
Qy 1312 TGGAAACAAAGTACTTGTGAATTTGTAAGGACCTTAATATGGAACACAAACAGGAGCTTCC 1371
Db 1480 TGGAAACAAAGTACTTGTGAATTTGTAAGGACCTTAATATGGAACACAAACAGGAGCTTCC 1539
Qy 1372 TGTGTCAGTACAGGACAGGCTGTGATGAGTGTGCTCATCTGACAAACATATCAGATTA 1431
Db 1540 TGTGTCAGTACAGGACAGGCTGTGATGAGTGTGCTCATCTGACAAACATATCAGGCTG 1599
Qy 1432 TGGGACATAGAAATGTGTCATTTTACAGTGTGTAAGGACCATCAGAAATTTGGTGGCT 1491
Db 1600 TGGGACATAGAAATGTGTCATTTTACAGTGTGTAAGGACCATCAGAAATTTGGTGGCT 1659
Qy 1492 TGTATTCGATTTGATTAACAGAGGATAGTCACTGAGGAGCTATGATGGGAAATTAAGTG 1551
Db 1660 TGCATTCGATTTGATTAACAGAGGATAGTCACTGAGGAGCTATGATGGGAAATTAAGTG 1719
Qy 1552 TGGGATCTTGTGGCTGTGTTGAGACCCCGCTGTCTCTGACGGGACACTCTGTCTACGGACC 1611
Db 1720 TGGGATCTTGTGGCTGTGTTGAGACCCCGCTGTCTCTGACGGGACACTCTGTCTACGGACC 1779
Qy 1612 CTTGTGG 1618
Db 1780 CTTGTGG 1786

```

RESULT 4

AK041532
LOCUS Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630019L11 product:F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2) homolog (Homo sapiens), full insert sequence.
ACCESSION AK041532
VERSION AK041532.1 GI:26334548
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)

ORGANISM	Mus musculus	TITLE	Direct Submission
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
AUTHORS	Carninci, P. and Hayashizaki, Y.	COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
TITLE	High-efficiency full-length cDNA cloning		Tissues were provided by Dr. John Todd (Dept. of Medical Genetics, Wellcome Trust Centre for Molecular Mechanisms in Disease, Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		Please visit our web site for further details.
MEDLINE	99279253		URL: http://genome.gsc.riken.go.jp/
PUBMED	10349636		URL: http://fantom.gsc.riken.go.jp/
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	FEATURES	Location/Qualifiers
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	source	1. .3824
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		/organism="Mus musculus"
MEDLINE	20499374		/mol_type="mRNA"
PUBMED	11042159		/strain="C57BL/6J"
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		/db_xref="FANTOM DB:A630019L11"
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplex capillary sequencer		/db_xref="taxon:10090"
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		/clone="A630019L11"
MEDLINE	20530913		/tissue_type="thymus"
PUBMED	11076861		/clone_lib="RIKEN full-length enriched mouse cDNA library"
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Rong, C., Quackenbush, J., Schriml, L.M., Stauble, F., Suzuki, G., Tomita, M., Walder, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Bagheri, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.P., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hall, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.	CDS	/codon_start=1
TITLE	Functional annotation of a full-length mouse cDNA collection		/protein_id="BAC30375.1"
JOURNAL	Nature 409 (6821), 685-690 (2001)		/db_xref="GI:26334549"
MEDLINE	21085660		/translations="MEPDSVIEDKTELMTNTSMVEDONEDSPKSLWQISNGTSSV IVSRKPSNGYOKEDLCIKYEDWSDQVEFVHLISRMCHVOGHJHNSYKPKML QREPTALPEQGLDHAENLNSYLDARSALCAELVCKEQRVISEGMWKLIERMVR TELWGLSGRRGNDQILFNSPTDGPNSFYRSLPKLIQDIETIESNRCGRHNLQ RIORSENSGVYCLQYDDDKIISLRDINSIKWDSKLECLKVLITGHTSVCLQYD ASATDITLRLVGHRAAVNVDFDDKYIVSAGDRTIKWMTSTCEFRVTLNGHRG IACQVDRDLVSGSSDNTIRLWDIECGACLVLEGEELVRCIFRPNKRVISGAYDG KIKVWLQALDPRAPASTLCRLTLVHSRVRFLQFDFEQLIISSHDDTLILWDFLN VPSPAQNETRSPSTYTYISR"
PUBMED	11217851		/notes="unnamed protein product; F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2) homolog [Homo sapiens] (SWISSPROT Q9UK81, evidence: PASTY, 99.2%ID, 100%length, match=1587) putative"
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		/dev_stage="3 days neonate"
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		169. .1758
JOURNAL	Nature 420, 563-573 (2002)		/notes="F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2) homolog [Homo sapiens] (SWISSPROT Q9UK81, evidence: PASTY, 99.2%ID, 100%length, match=1587) putative"
PUBMED	1217851		/protein_id="BAC30375.1"
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	polyA_signal	3805. .3810
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	polyA_site	3824
JOURNAL	Nature 420, 563-573 (2002)		/notes="putative"
PUBMED	1217851		/notes="putative"
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		BASE COUNT 885 a 915 c 982 g 1042 t
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		ORIGIN
JOURNAL	Nature 420, 563-573 (2002)		Query Match 40.8%; Score 877.8; DB 11; Length 3824;
PUBMED	1217851		Best Local Similarity 74.3%; Pred. No. 4.3e-233;
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		Matches 1120; Conservative 0; Mismatches 377; Indels 6; Gaps 1;
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		QY 298 AATGGCACTTCCAGTATGATTTGTGCCCAACGACGAAACTCTCAGCAAGCTATGAAAAG 357
JOURNAL	Nature 420, 563-573 (2002)		Db 283 AATGGAACGTCATCTGTGATTTGTTCTCCAGAAAGCCGTCAGAGGGGACTACCGAAA 342
PUBMED	1217851		QY 358 GAAAGAACTGTGTCTGTAATACTTTTGAGCGAGTGGTCTAGAGTCAGATCAAGTGAATTT 417
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		Db 343 GAAAGAACTGTGTGATTAAGTACTTTTGACAGTGGTCTGTAATCAGATCAGGTGAATTT 402
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		QY 418 GTGGAACATCTTATATCCCAAAATGTCTCATTTACCAACATGGGCACATTAACCTCGTATCTT 477
JOURNAL	Nature 420, 563-573 (2002)		Db 403 GTGGAGCATCTTATCTCAAGGATGTCTCATTTATCAGATGGACATTAACATCTTACCTG 462


```
256 GCGTGTGATTCGATTTGATTAACAGAGGATAGTCAGTGGGCGCTATGATGGAATAATTAA 315
1548 AGTGTGGGATCTTTGTGGCTGCTTTGGACCCCGTCTCTCCAGGGACACTCTGTCTACG 1607
316 AGTGTGGGATCTTTGTGGCTGCTTTGGACCCCGTCTCTCCAGGGACACTCTGTCTACG 375
1608 GACCTTGTGGAGCATTCGGGAAGATTTTTCGACTACAGTTGATGAATCCAGATTGT 1667
376 GACCTTGTGGAGCATTCGGGAAGATTTTTCGACTACAGTTGATGAATCCAGATTGT 435
1668 CAGTAGTTCACATGATGACACAAATCTCATCTGGGACTTCTAAATGATCCAGTGGCCA 1727
436 CAGTAGTTCACATGATGACACAAATCTCATCTGGGACTTCTAAATGATCCAGTGGCCA 495
1728 AGCTGAACCCCGGCTTCCCTTTCGAAACATACACTACATCTCCAGATAAATAACCAT 1787
496 AGCTGAACCCCGGCTTCCCTTTCGAAACATACACTACATCTCCAGATAAATAACCAT 555
1788 ACAGTCACTCATACTTGGCCAGGACCAATTAAGTTGCGGTATTTAAAGTATCTGCCAA 1847
556 ACAGTCACTCATACTTGGCCAGGACCAATTAAGTTGCGGTATTTAAAGTATCTGCCAA 615
1848 TACCAGGATGACACACAGTAACAACTACTGCCCCAGTTTCCCTGGACTAGCG 1907
616 TACCAGGATGACACACAGTAACAACTACTGCCCCAGTTTCCCTGGACTAGCG 675
1908 AGGAGCAGGGCT-TTGAGACTCTCTTGGGACACAGTTGCTCAGTTCGCGCCAGGACG 1966
676 AGGAGCAGGGCTATTGAGACTCTCTTGGGACACAGTTGCTCAGTTCGCGCCAGGACG 735
1967 GTCTACTAGCACTGACTGCTTCAAGTGTGCTATCAGAAGATGCTTCTATCAATTG 2026
736 GTCTACTAGCACTGACTGCTTCAAGTGTGCTATCAGAAGATGCTTCTATCAATTG 795
2027 TGAATGATTGGAACCTTTAAACCTCCCTCTCTCCCTCTTCACTCTGACCTAGTTT 2086
796 TGAATGATTGGAACCTTTAAACCTCCCTCTCTCCCTCTTCACTCTGACCTAGTTT 855
2087 TTTCCATTGTTCCAGACAAAGTGACTATAAATAATTAGTTGTTTCCAGAA 2143
856 TTTTCCATTGTTCCAGACAAAGTGACTATAAATAATTAGTTTCCAGAA 912
```

```
RESULT 6
LOCUS BE782628
DEFINITION 601465751F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3868812 5',
mRNA sequence.
ACCESSION BE782628
VERSION BE782628.1 GI:10203826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9617 row: f column: 13
High quality sequence stop: 655.
FEATURES
Location/Qualifiers
1..709
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3868812"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: Noif;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
```

BASE COUNT 193 a 156 c 175 g 185 t
ORIGIN

Query Match 30.2%; Score 650.6; DB 10; Length 709;
Best Local Similarity 99.2%; Pred. No. 4.9e-170;
Matches 700; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

```
QY 1222 GTGCTGTGCGGACACCGAGCTGCTGTCATGTTGTAGACTTTGATGACAAAGTACATTGTT 1281
DB 1 GTGCTGTGCGGACACCGAGCTGCTGTCATGTTGTAGACTTTGATGACAAAGTACATA-TGTT 59
QY 1282 TCTGCATCTGGGATAGAACTATAAAGGTATGAAACACAAGTACTTGTGAATTTGTAAGG 1341
DB 60 TCTGCATCTGGGATAGAACTATAAAGGTATGAAACACAAGTACTTGTGAATTTGTAAGG 119
QY 1342 ACCTTAAATGGACACAAACGAGGCATTGCTGTTTCAGTACAGGACAGGCTGCTAGTG 1401
DB 120 ACCTTAAATGGACACAAACGAGGCATTGCTGTTTCAGTACAGGACAGGCTGCTAGTG 179
QY 1402 AGTGGCTCATCTGACACACTATCAGATTATGGACATAGAAATGTGTCATGTTTACGA 1461
DB 180 AGTGGCTCATCTGACACACTATCAGATTATGGACATAGAAATGTGTCATGTTTACGA 239
QY 1462 GTGTTAAGAGCCATGAGGATTTGGTGGCTGTTGATTCGATTTGATTAACAAGAGGATGTC 1521
DB 240 GTGTTAAGAGCCATGAGGATTTGGTGGCTGTTGATTCGATTTGATTAACAAGAGGATGTC 299
QY 1522 AGTGGGCTCATGATGGAATAAAGTGTGGATCTTTGTGGCTGCTTTGGACCCCGT 1581
DB 300 AGTGGGCTCATGATGGAATAAAGTGTGGATCTTTGTGGCTGCTTTGGACCCCGT 359
QY 1582 GCTCCTCGAGGACACTCTGCTACGAGCCCTGTGGACATTCGGAAAGATTTTGA 1641
DB 360 GCTCCTCGAGGACACTCTGCTACGAGCCCTGTGGACATTCGGAAAGATTTTGA 419
QY 1642 CTACAGTTTGTGATGAATTCAGATTGTGAGTAGTTTCATGATGACACAAATCCTCATCTGG 1701
DB 420 CTACAGTTTGTGATGAATTCAGATTGTGAGTAGTTTCATGATGACACAAATCCTCATCTGG 479
QY 1702 GACTTCTTAATGATCCAGTGCCTGAGCTGAACCCCGGTTCCCTTCTCGAATACATAC 1761
DB 480 GACTTCTTAATGATCCAGTGCCTGAGCTGAACCCCGGTTCCCTTCTCGAATACATAC 538
QY 1762 ACCTACATCTCCAGATAAATAACCATACACTGACCTCATCTATTTGCCAGGCCATTA 1821
DB 539 ACCTACATCTCCAGATAAATAACCATACACTGACCTCATCTATTTGCCAGGCCATTA 598
QY 1822 GTTCCGTTATTAACGATATCTGCAATACAGATGAGCAACAAAGTAACTCAACT 1881
DB 599 GTTCCGTTA-TTAACTGATCTGCAATACAGATGAGCAACAAAGTAACTCAACT-NACT 656
QY 1882 ACTGCCAGTTTCCCTGGACTAGCCAGGAGGAGGCTTTGAGACTCCTGTG 1934
DB 657 AGTGGCCAGTTTCCCTGGACTAGCCAGGAGGAGGCTTTGAGACTCCTGTG 709
```

RESULT 7
LOCUS CB244526
DEFINITION UI-M-FYO-cdg-k-22-0-UT, ri NIH_BMAP_FYO Mus musculus cDNA clone
IMAGE: 683335 5', mRNA sequence.
ACCESSION CB244526
VERSION CB244526.1 GI:28366170

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BNAP). "Gene Discovery in the Developing Mouse Nervous System", supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 223 a 181 c 204 g 169 t 1 others
ORIGIN

Query Match 29.7%; Score 638; DB 13; Length 778;
Best Local Similarity 89.5%; Pred. No. 1.7e-166; Indels 1; Gaps 1;
Matches 697; Conservative 0; Mismatches 81

256 GCTATGAAGACTGAGAAATTTGTGGCCAAACAAACAACTTGCATATGGCACTTCAGATG 315
|||||
257 GCTATGAAGACTGAGAAATTTGTGGCCAAACAAACAACTTGCATATGGCACTTCAGATG 315
|||||

316 ATTGTGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 375
|||||

61 ATTGTGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 420
|||||

376 AATATCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTGGAACTTTTATATCC 435
|||||

121 AAGTATTTTGGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTAGAACACCTTATATCC 180
|||||

436 CAATGTGTCATACCAATGGGCAATTAACCTGATCTTAACTATGTCGACAGA 495
|||||

181 CAATGTGTCATACCAATGGGCAATTAACCTGATCTTAACTATGTCGACAGA 240
|||||

496 GATTTCAATCACTGCTGCGCAGTGGGGAATGGATCATATCGTGGAAACATTTCTGTCA 555
|||||

241 GATTTCAATCACTGCTGCGCAGTGGGGAATGGATCATATCGTGGAAACATTTCTGTCA 300
|||||

556 TACTGTGATGCGCAATCACTATGCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 615
|||||

301 TACTGTGATGCGCAATCACTATGCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 360
|||||

616 ACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGATGGTTCAGGACAGATTTCTCTG 675
|||||

361 ACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGATGGTTCAGGACAGATTTCTCTG 420
|||||

676 TGGAGAGGCTGCGCAAGCAAGAGAGATGGGCAAGATTTATTCAGAAACAACTTCCT 735
|||||

421 TGGCAGAGGCTGCGCAAGCAAGAGAGATGGGCAAGATTTATTCAGAAACAACTTCCT 480
|||||

736 GAGCGGAATGCTCTCTCCCACTCTTTTATAGAGCACTTTATCTCTAAATTTATACAGAC 795
|||||

481 GATGAGAAAGCTCTCTCCCACTCTTTTATAGAGCACTTTATCTCTAAATTTATACAGAC 540
|||||

796 ATTGAGACATAGATCTTAATTTGAGATGTGAAGACATAGTTTACAGAGATTTACATGC 855
|||||

541 ATTGAGACATAGATCTTAATTTGAGATGTGAAGACATAGTTTACAGAGATTTACATGC 600
|||||

856 CGAAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGATCGAGAAATAGTAAGC 915
|||||

601 CGAAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGATCGAGAAATAGTAAGC 660
|||||

916 GGCCTTCGACACACACATCAGATCTGGGATTAACACATTTGGATGTGAACGGCAATT 975
|||||

661 GGCCTTCGACACACACATCAGATCTGGGATTAACACATTTGGATGTGAACGGCAATT 719
|||||

976 CTCACAGGCATACAGGTTTCAGTCTCTCTCCAGTATGATGAGAGAGTGCATATAC 1034
|||||

720 CTCACAGGCATACAGGTTTCAGTCTCTCTCCAGTATGATGAGAGAGTGCATATAC 778
|||||

RESULT 9
LOCUS AL044095
DEFINITION DKFZp34M1528_r1 434 (synonym: htes3) Homo sapiens cDNA clone
AL044095
ACCESSION AL044095
VERSION AL044095.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 652)
Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Bloecker, et al.)
Unpublished
Contact: Bloecker H
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
sequenced by GPF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp34M1528) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..652
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp34M1528"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 165 a 166 c 138 g 183 t
ORIGIN

Query Match 29.5%; Score 635.2; DB 9; Length 652;
Best Local Similarity 99.2%; Pred. No. 9.3e-166;
Matches 649; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1479 GGAATTCGTGGTGTGATTCGATTTGATTAACAGAGATAGTCAGTGGGCGCTATGATGG 1538
Db 1 GGAATTCGTGGTGTGATTCGATTTGATTAACAGAGATAGTCAGTGGGCGCTATGATGG 60
QY 1539 AAAAATTAAGTGTGGATCTTGTGGCTGCTTTGGACCCCGCTGCTTCGAGGACACT 1598
Db 61 AAAAATTAAGTGTGGATCTTGTGGCTGCTTTGGACCCCGCTGCTTCGAGGACACT 120
QY 1599 CTGCTTAGGACCTTGTGGAGCATTCGGAGAGTTCGACTACAGTTTGATTAATT 1658
Db 121 CTGCTTAGGACCTTGTGGAGCATTCGGAGAGTTCGACTACAGTTTGATTAATT 180
QY 1659 CCAGATTGTTCAGTAGTTTCATGATGACAAATCTCTCTCTGGGACTTCCTAAATGATCC 1718
Db 181 CCAGATTGTTCAGTAGTTTCATGATGACAAATC--CATCTGGACTTCCTAAATGATCC 238
QY 1719 AGCTGCCCAAGCTGAACCCCGCTTCCCTTCTCGAATACATACCTATCTCCAGATA 1778
Db 239 AGCTGCCCAAGCTGAACCCCGCTTCCCTTCTCGAATACATACCTATCTCCAGATA 298
QY 1779 AATAACCATACACTGACCTCATCTTGTCCCGAGGACCCATTAAGTTGCGGTATTTAACT 1838
Db 299 AATAACCATACACTGACCTCATCTTGTCCCGAGGACCCATTAAGTTGCGGTATTTAACT 358

```

1839 ATCTGCAATACAGATGAGCAACACAGTAACAATCAAACTACTGCCAGTTTCCCTG 1898
1899 GACTAGCCGAGGAGCAGGCTTTGAGACTCCGTTGGGACACAGTTGCTGCGAGTCGC 1958
419 GACTAGCCGAGGAGCAGGCTTTGAGACTCCGTTGGGACACAGTTGCTGCGAGTCGC 478
1959 CCAGGACGCTTACTCAGCACAACTGACTGCTTCAGTGTGCTATCAGAGATGCTTCT 2018
479 CCAGGACGCTTACTCAGCACAACTGACTGCTTCAGTGTGCTATCAGAGATGCTTCT 538
2019 ATCAATGTGAATGATGAGACTTTTAAACCTCCCTCTCCCTCCCTTTCACCTCTGCA 2078
539 ATCTTTGTGAATGATGAGACTTTTAAACCTCCCTCTCCCTCTCCCTTTCACCTCTGCA 598
2079 CCTAGTTTTTCCATGTTTCCAGACAAAGTGACTTATAATATATATTAGTG 2132
599 CCTAGTTTTTCCATGTTTCCAGACAAAGTGACTTATAATATATATTAGTG 652

RESULT 10
BQ425358
LOCUS
DEFINITION
AGENCOURT_7907367 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154838
5', mRNA sequence.
ACCESSION
BQ425358.1 GI:21120673
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 851)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13496 row: i column: 15
High quality sequence stop: 697.

FEATURES
source
Location/Qualifiers
1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6154838"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-Sport6; Site 1: Not I;
Site 2: Sal I; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 243 a 196 c 223 g 188 t 1 others
ORIGIN
Query Match 29.1%; Score 625; DB 13; Length 851;
Best Local Similarity 99.1%; Pred. No. 7.6e-163;
Matches 639; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

112 TTTATGAATTCCTCAGAGAGAGAGACTGTATATATGGCAACCCCTAGGAAGATAATA 171
207 TTCCAGAATTCCTCAGAGAGAGAGACTGTATATATGGCAACCCCTAGGAAGATAATA 266

```

```

QY 1172 CCAGAGAGAGAAATTCACCTTAGACAGACATACAAAGCTGTGCCAGACTCTGCTTAAACCAA 231
DB 267 CCAGAGAGAGAAATTCACCTTAGACAGACATACAAAGCTGTGCCAGACTCTGCTTAAACCAA 326
QY 232 GAAACAGTATGTTTATAGCAAGCACTGTATGAAGCTGAGAAATTTGTGTGCCCAAAACAAA 291
DB 327 GAAACAGTATGTTTATAGCAAGCACTGTATGAAGCTGAGAAATTTGTGTGCCCAAAACAAA 386
QY 232 CTTGCCAATGCCACTTCCAGTATGATTTGCCCAAGCAACGAAACTCTCAGCAAGCTAT 351
DB 387 CTTGCCAATGCCACTTCCAGTATGATTTGCCCAAGCAACGAAACTCTCAGCAAGCTAT 446
QY 352 GAAAGAGGAAAGAGAACTGTGTCAAAATCTTTGAGCAGTGTGTCAGATCAGATCAAGTG 411
DB 447 GAAAGAGGAAAGAGAACTGTGTCAAAATCTTTGAGCAGTGTGTCAGATCAGATCAAGTG 506
QY 412 GAATTTGTGGAACTTCATATATATATATATATATATATATATATATATATATATATAT 471
DB 507 GAATTTGTGGAACTTCATATATATATATATATATATATATATATATATATATATATAT 566
QY 472 TATCTTAAACCTATGTTGCAGAGAGATTTCAATCACTGCTGCCAGCTGGGATTTGGAT 531
DB 567 TATCTTAAACCTATGTTGCAGAGAGATTTCAATCACTGCTGCCAGCTGGGATTTGGAT 626
QY 532 CATATCGCTGAGAACTTCTGTATACCTGGATGCCAAATCACTATGCTGCTGAATTT 591
DB 627 CATATCGCTGAGAACTTCTGTATACCTGGATGCCAAATCACTATGCTGCTGAATTT 686
QY 592 GTGTGCRAGGAATGTTACCGAGTGACCTCTGTAGTGAGTCTGTGGAGAGAGCTTATCGAG 651
DB 687 GTGTGCRAGGAATGTTACCGAGTGACCTCTGTAGTGAGTCTGTGGAGAGAGCTTATCGAG 746
QY 652 AGAATGTCAGGACAGATTTCTGTGGAGAGGCTGTGCAGAACGAGAGATGGGG-ACA 710
DB 747 AGAATGTCAGGACAGATTTCTGTGGAGAGGCTGTGCAGAACGAGAGATGGGGNACA 806
QY 711 GTATTTATTTCAAAAACAAACCTCTCTGACGGGAATGCTCTCTCCCA 755
DB 807 GTATTTATTTCAAAAACAAACCTCTCTGACGGGAATGCTCTCTCCCA 851

RESULT 11
BQ465280
LOCUS
DEFINITION
AGENCOURT_6427660 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502075
5', mRNA sequence.
ACCESSION
BQ465280
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 973)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12139 row: k column: 04
High quality sequence stop: 601.

FEATURES
source
Location/Qualifiers
1..973
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```
/clone="IMAGE:5502075"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_57"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT      282 a      245 g      246 t
ORIGIN
Query Match      29.0%; Score 624.8; DB 12; Length 973;
Best Local Similarity 97.2%; Pred. No. 9.3e-163;
Matches 690; Conservative 0; Mismatches 12; Indels 8; Gaps 5;
2Y 658 GTCCGACAGATCTCTGTGGAGAGGCTCGCAGAACGAGAGGATGGGACAGTATTTA 717
2b 20 GTCCGACAGATCTCTGTGGAGAGGCTCGCAGAACGAGAGGATGGGACAGTATTTA 79
2Y 718 TTCAAAAACAAACCTCTGAGGGAGTCTCTCCCACTCTTTTATAGACACTTAT 777
2b 80 TTCAAAAACAAACCTCTGAGGGAGTCTCTCCCACTCTTTTATAGACACTTAT 139
2Y 778 CTTAAAAATTATCAAGACATTTGAGCAATAGACTTAATTGGAGATGTGAAGACATAGT 837
2b 140 CTTAAAAATTATCAAGACATTTGAGCAATAGACTTAATTGGAGATGTGAAGACATAGT 199
2Y 838 TTACAGAGATTTCACTGCGGAAGTGAAACAGCAAGAGGTTTACTGTTTACAGTATGAT 897
2b 200 TTACAGAGATTTCACTGCGGAAGTGAAACAGCAAGAGGTTTACTGTTTACAGTATGAT 259
2Y 898 GATCGAAAAATAGTAAGCGGCTTCGAGACAAACAATCAAGATCTCGGATAAAAAACA 957
2b 260 GATCGAAAAATAGTAAGCGGCTTCGAGACAAACAATCAAGATCTCGGATAAAAAACA 319
2Y 958 TTGGATGCAAGCGATTTCTACAGGCCATACAGGTCAGTCTCTCTCTCCAGTATGAT 1017
2b 320 TTGGATGCAAGCGATTTCTACAGGCCATACAGGTCAGTCTCTCTCTCCAGTATGAT 379
2Y 1018 GAGAGAGTGATCATAAACAGGATTCAGGATTCACGGTTCAGAGTGTGGGATGTAATAACA 1077
2b 380 GAGAGAGTGATCATAAACAGGATTCAGGATTCACGGTTCAGAGTGTGGGATGTAATAACA 439
2Y 1078 GGTGAATGCTTAAACAGTGTGATTCACCATTTGTGAAGAGTTCTGCACTTGGCTTTCAAT 1137
2b 440 GGTGAATGCTTAAACAGTGTGATTCACCATTTGTGAAGAGTTCTGCACTTGGCTTTCAAT 499
2Y 1138 AATGGCATGATGGTGACCTGCTCCAAAGATCGTTCCCATTTGTGTATGGGATATGGCTCC 1197
2b 500 AATGGCATGATGGTGACCTGCTCCAAAGATCGTTCCCATTTGTGTATGGGATATGGCTCC 559
2Y 1198 CCAACTGACATTTACCTCTGGGAGGTGCTGTGGACACC-GAGCTGCTCAATGTTGT 1256
2b 560 CCAACTGACATTTACCTCTGGGAGGTGCTGTGGACACC-GAGCTGCTCAATGTTGT 619
2Y 1257 AGACTTTGATGACAAAGTACATTG-TTTCTGCATCTGGGG--ATAGAATCTATAAGGTATG 1313
2b 620 AGACTTTGATGACAAAGTACATTGTTTCTGCATCTGGGGATAGAACTATAAGGTATG 679
2Y 1314 GAACA-CAAGTACTTGTGAATTTGT---AAGACCTTTAAATGGACACAAA 1359
2b 680 GAACACCAAGTACTTGTGAATTTTGTGAAGACCTTTAAATGGACACAAA 729
RESULT 12
BUS07728
LOCUS      BUS07728      887 bp      mRNA      linear      EST 12-SEP-2002
DEFINITION AGENCOURT_10095724 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6501709
5', mRNA sequence.
ACCESSION  BUS07728
VERSION    BUS07728.1 GI:22813961
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
```

```
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 887)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgabs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LHAM14056 row: j column: 14
              High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source           1..887
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6501709"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                     Average insert size 2.1 kb."
BASE COUNT      218 a      239 c      200 g      228 t
ORIGIN
Query Match      28.9%; Score 622.2; DB 13; Length 887;
Best Local Similarity 98.7%; Pred. No. 4.7e-162;
Matches 627; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1509 CAAGAGGATAGTCAGTGGGGCCCTATGATGGAATAAATGATGCGGATCTTGTGGTGC 61
2b 2 CACGCGTCCGTCAGTGGGGCCCTATGATGGAATAAATGATGCGGATCTTGTGGTGC 61
QY 1569 TTTGGACCCCGTCTCTGAGGACACTCTGTACGGACCTTGTGGACCTTGGGACATTCGG 1628
2b 62 TTTGGACCCCGTCTCTGAGGACACTCTGTACGGACCTTGTGGACATTCGG 121
QY 1629 AAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTTGTCAGTAGTTCACATGATGAC 1688
2b 122 AAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTTGTCAGTAGTTCACATGATGAC 181
QY 1689 AATCCTCATCTGGGACTTCTTAATGATCCAGTCCGCAAGCTGAACCCCGTTCGCC 1748
2b 182 AATCCTCATCTGGGACTTCTTAATGATCCAGTCCGCAAGCTGAACCCCGTTCGCC 241
QY 1749 TTCTCGAACAATACACCTTACATCTCCAGATAAATAACCATACACTGACCTCATCTTGGCC 1808
2b 242 TTCTCGAACAATACACCTTACATCTCCAGATAAATAACCATACACTGACCTCATCTTGGCC 301
QY 1809 AGNACCCATTAAGTTGGGTATTTAACTATCTCCCAATACAGGATGAGCAACAG 1868
2b 302 AGNACCCATTAAGTTGGGTATTTAACTATCTCCCAATACAGGATGAGCAACAG 361
QY 1869 TAACAATCAAACTACTGCCCAGTTTCCCTGGACTAGCCAGGAGGAGGCTTTGAGACTC 1928
2b 362 TAACAATCAAACTACTGCCCAGTTTCCCTGGACTAGCCAGGAGGAGGCTTTGAGACTC 421
QY 1929 CTGTTGGGACACAGTTGGTCTGAGTGGCCCGGACGAGGCTTACTCAGCACTGACTG 1988
2b 422 CTGTTGGGACACAGTTGGTCTGAGTGGCCCGGACGAGGCTTACTCAGCACTGACTG 481
QY 1989 CTTTCAGTGTCTATCAGAAAGATGCTTCTATCAATTTGTAATGATTGGAATTTTAAAC 2048
2b 482 CTTTCAGTGTCTATCAGAAAGATGCTTCTATCAATTTGTAATGATTGGAATTTTAAAC 541
```

2Y 2049 CTCCCTCTCTCCCTCTTACCTCTGCACTAGTCTTTTCCCATGTTCCAGACAA 2108
 Db 542 CTCCCTCTCTCTCCCTCTTACCTCTGCACTAGTCTTTTCCCATGTTCCAGACAA 601
 2Y 2109 GGTGACTTATAATATATTAGTGTGTTTCCAGAA 2143
 Db 602 GGTGACTTATAATATATTAGTGTGTTTCCAGAA 636

RESULT 13
 CB520936 744 bp mRNA linear EST 28-MAR-2003
 LOCUS UI-M-GH0-cel-p-18-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
 DEFINITION IMAGE: 6841315 5', mRNA sequence.
 CB520936
 CB520936.1 GI:29354291
 EST.
 Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 744)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5,
 Location/Qualifiers
 1..744
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6841315"
 /tissue_type="Whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP GH0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaïdo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CCAATGAAAT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

BASE COUNT 197 a 184 c 202 g 160 t 1 others
 ORIGIN
 Query Match 28.7%; Score 616.6; DB 14; Length 744;
 Best Local Similarity 89.2%; Pred. No. 1.6e-160;
 Matches 664; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

2Y 501. CATACTGCTCGCCAGCTCGGGGATTGGATCATATCGCTGAGAACATTTGTCATACCT 560
 Db 1 CATACTGCTCGCCAGCAGCGGGGCTCGGACCACATCGCTGAGAACATTTGTCATACCT 60

QY 561 GGATGCCAATCACTATGTGCTGCTGAATCTTGTCGAGGAATGTAACCGAGTCACTC 620
 Db 61 GGAGCCCAAGTCACTGTGCTGCTGAGCTCGTGTGCAAGGAATGGTACCCTGACCTC 120
 QY 621 TGATGGCATCTCTGGAAGAAGCTTATCGAGAGAATGGTCAGGACAGATTCTCTGTGAG 680
 Db 121 GGACGGCATCTCTGGAAAAAGCTCATCGAGAGGATGGTCAGGACGGACTCTCTGTGCG 180
 QY 681 AGGCCTGGCAGACGAGAGGATGGGACAGTATTATTCAAAAACAACCTCTCTGAGG 740
 Db 181 AGGCCTGGCAGAGCGCAGAGGCTGGGGACAGTACTATTTCAAAAACAACCTCTCTGATGA 240
 QY 741 GAATGCTCTCTCCCAACTCTTTTATAGACACTTTTATCCTAAAAATTATACAGACATTGA 800
 Db 241 GACGCTCTCTCCCAACTCTTTTATAGAGCGCTTTATCTCTAAATCATACAGACATTGA 300
 QY 801 GACATAGATCTTAATGGAGATGGGAAGACATAGTTTACAGAAATCACTGCCGAAG 860
 Db 301 GACAAATAGATCTCAATTGGAGATGGGGCGACATAGTTTACAGAGAAATCCACTGCCGGAG 360
 QY 861 TGAACAAGCAAAAGGAGTTTACTGTTTACAGTATGATGATCAGAAAAATAGTAAGCGGCT 920
 Db 361 TGAACAAGTAAAGGGTTTACTGTTTACAGTACGACGACGAGAGATAGTCAGCGGCT 420
 QY 921 TCGAGACAAACAAATCAAGATCTGGATTAACACATTTGGAATCGAAGCGAATTTCTCAC 980
 Db 421 TCGAGACAAACAAATCAAGATCTGGATTAACACATTTGGAATCGAAGCGAATTTCTCAC 480
 QY 981 AGGCATACAGGTTTCAGTCTCTGTCCTCCAGTATGATGAGAGAGTGATCATAACAGGATC 1040
 Db 481 GGGCCACACGGGCTCGTCTGTCCTGTCAGTACGATGAGAGGGTGATCATCAGAGGCTC 540
 QY 1041 ATCGATTCACCGGTGAGAGTGGGATGTAATACAGTGAAATGCTAAACACGCTTGAT 1100
 Db 541 CTCAGATCCACCGTCCAGAGTGTGGATGTAATGAGTGTGAGATGCTAAACACATTTGAT 600
 QY 1101 TCACCATTTGAAGCAGTTCTGCACTTGCGTTTCAATATGTCATGATGGTGCACCTGCTC 1160
 Db 601 TCACCATTTGAAGCAGTTCTGCACTTGCGTTTCAATATGTCATGATGGTGCACCTGCTC 660
 QY 1161 CAAGATCGTTCATGCTGATGGGATATGGCTCCCAACTGATACCTCCCGGAG 1220
 Db 661 CAAGACCGTTCATGCTGATGGGATATGGCTCCCAACTGATACCTCCCGGAG 720
 QY 1221 GGTGCTGTCGGACACCGAGCTGC 1244
 Db 721 GGTGCTGTCGGACACCGAGCTGC 744

BE296484 758 bp mRNA linear EST 20-JUL-2000
 601174756F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3530341 5',
 mRNA sequence.
 BE296484
 BE296484.1 GI:9180046
 EST.
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 758)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLM202 row: o column: 14
High quality sequence stop: 617.

```

FEATURES
  source
    1..758
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:3530341"
      /tissue_type="thabdomyosarcoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH MGC 17"
      /notes="Organ: muscle; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
  BASE COUNT      212 a 149 c 202 g 195 t
  ORIGIN
    Query Match      28.5%; Score 613.2; DB 10; Length 758;
    Best Local Similarity 96.3%; Pred. No. 1.4e-159;
    Matches 649; Conservative 0; Mismatches 23; Indels 2; Gaps 2;
  2y 602 AATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGATGGTCA 661
  Db 1 AATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGATGGTCA 60
  2y 662 GGACAGATTTCTGTGGAGAGCGCTGGGAGACGAGAGAGATGGGACAGATTATTCA 721
  Db 61 GGACAGATTTCTGTGGAGAGCGCTGGGAGACGAGAGAGATGGGACAGATTATTCA 120
  2y 722 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 781
  Db 121 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 180
  2y 782 AAATTATACAGACATTTGAGACAATAGAACTCTAATTGGAGATGTGGAGACATAGTTTAC 841
  Db 181 AAATTATACAGACATTTGAGACAATAGAACTCTAATTGGAGATGTGGAGACATAGTTTAC 240
  2y 842 AGAGATTCTACTCCCAAGTGAACAGCAAGAGAGTTTACTGTTACAGTATGATGATC 901
  Db 241 AGAGATTCTACTCCCAAGTGAACAGCAAGAGAGTTTACTGTTACAGTATGATGATC 300
  2y 902 AGAAATAGTAAAGCGGCTTCGAGACACACAACTCAAGATCTGGATTAACACACATTGG 961
  Db 301 AGAAATAGTAAAGCGGCTTCGAGACACACAACTCAAGATCTGGATTAACACACATTGG 360
  2y 962 AATGCGAGGAATTTCTCAGGCGCATACAGGTTTCAGTCTCTGTCTCCAGTATGATGAGA 1021
  Db 361 AATGCGAGGAATTTCTCAGGCGCATACAGGTTTCAGTCTCTGTCTCCAGTATGATGAGA 420
  2y 1022 GAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGTGGAGTGT-AAAACACAGGT 1080
  Db 421 GAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGTGGAGTGT-AAAACACAGGT 480
  2y 1081 GAAATGCTAAACACGTTGATTACCAATTTGAGACAGTTCTGCACCTTCGCTTTCAATAAT 1140
  Db 481 GAAATGCTAAACACGTTGATTACCAATTTGAGACAGTTCTGCACCTTCGCTTTCAATAAT 540
  2y 1141 GGCATGATGTGACCTGTCTCAAAGATCGTTCCATTTGCTGTATGGGATATGGCCTCCCCA 1200
  Db 541 GGCATGATGTGACCTGTCTCAAAGATCGTTCCATTTGCTGTATGGGATATGGCCTCCCCA 600
  2y 1201 ACTGACATTTACCTTCGGAGGGTGTGGTC-GGACACGAGCTGCTGCATGTGTGTAGA 1259
  Db 601 CTGACATTTACCTTCGGAGGGTGTGGTCGGGACCCACGAGCTCTGCTCATGTGTGTGA 660
  2y 1260 CTTTGTGACCAAGT 1273
  Db 661 CTTTGTGTCGGT 674

```

```

RESULT 15
LOCUS BE269297
DEFINITION BE269297 Homo sapiens cdna clone IMAGE:3544549 5', mRNA sequence.
ACCESSION BE269297 GI:9142915
VERSION BE269297
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 692)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: CGapbs-remail.nih.gov
  Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
  cDNA Library Preparation: Ling Hong/Rubin Laboratory
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov.
  Plate: LLM239 row: o column: 14
  High quality sequence stop: 610.
  Location/Qualifiers
    1..692
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:3544549"
      /tissue_type="Burkitt lymphoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH MGC 8"
      /notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
  BASE COUNT      202 a 134 c 179 g 177 t
  ORIGIN
    Query Match      28.3%; Score 608.8; DB 10; Length 692;
    Best Local Similarity 97.7%; Pred. No. 2.3e-158;
    Matches 649; Conservative 0; Mismatches 12; Indels 3; Gaps 3;
  Qy 602 AATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGAAATGGTCA 661
  Db 1 AATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGAAATGGTCA 60
  Qy 662 GSACAGATTTCTCTGAGAGAGGCTGCGAAGAGAGAGATGGGACAGATTTATTCA 721
  Db 61 GSACAGATTTCTCTGAGAGAGGCTGCGAAGAGAGAGATGGGACAGATTTATTCA 120
  Qy 722 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 781
  Db 121 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 180
  Qy 782 AAATTATACAGACATTTGAGACAATAGAACTCTAATTGGAGATGTGGAGACATAGTTTAC 841
  Db 181 AAATTATACAGACATTTGAGACAATAGAACTCTAATTGGAGATGTGGAGACATAGTTTAC 240
  Qy 842 AGAGATTCTACTCCGAGGTGAACAGCAAGAGAGGTTTACTGTTTACAGTATGATGATC 901
  Db 241 AGAGATTCTACTCCGAGGTGAACAGCAAGAGAGGTTTACTGTTTACAGTATGATGATC 300
  Qy 902 AGAAATAGTAAAGCGGCTTCGAGACACACAACTCAAGATCTGGGATTAACACACATTGG 961

```


Db	301	AGAAATAGTAACGGCCTTCGAGACACACATCAAGATCTGGATTAACACACATTGG	360
Qy	962	AATGCAAGCGAATCTCAACAGGCCATACAGGTTCAAGTCTCTCTCTCCAGTATGATGAGA	1021
Db	361	AATGCAAGCGAATCTCAACAGGCCATACAGGTTCAAGTCTCTCTCTCCAGTATGATGAGA	420
Qy	1022	GAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTAATAACAGGTG	1081
Db	421	GAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTAATAACAGGTG	480
Qy	1082	AAATGCTAAACAGGTTGATTACCATTTGAAGCAGTTCTGCACCTTGGCTTCAATAATG	1141
Db	481	AAATGCTAAACAGGTTGATTACCATTTGAAGCAGTTCTGCACCTTGGCTTCAATAATG	540
Qy	1142	GCATGATGGTGACCT-GCTCCAAAGATCGTTCCATTGCTGTATGGATATGCGCTCCCCA	1200
Db	541	GCATGATGGTGACCTGGCTCCAAAGATCGTTCCATTGCTGTATGGATATGCGCTCCCCA	599
Qy	1201	ACTGACATTACCTCCGAGGGTGTGTGTCGACACCGAGCTGCTGTCATGTTGTAGAC	1260
Db	600	CTGACATTACCTCCGAGGGTGTGTGTCGACACCGAGCTGCTGTCATGTTGTAGAC	658
Qy	1261	TTTG 1264	
Db	659	TTGG 662	

Search completed: October 22, 2003, 21:57:19
 Job time : 4102 secs

QY	1621	CATTCCGGAAGAGTTTTCGACTACAGTTTGATGAATTCAGATTGTCAGTAGTTTCACAT	1680
DB	1621	CATTCCGGAAGAGTTTTCGACTACAGTTTGATGAATTCAGATTGTCAGTAGTTTCACAT	1680
QY	1681	GATGACAAATCCCTCATCTCGGAGCTTCCCTAAATGATGCAGCTGCCCAAGCTGAACCCCCC	1740
DB	1681	GATGACAAATCCCTCATCTCGGAGCTTCCCTAAATGATGCAGCTGCCCAAGCTGAACCCCCC	1740
QY	1741	CGTTCCCTTCTCGAACAATACACTACATCTCCAGATAAATAACCATACACTGCACCTCAT	1800
DB	1741	CGTTCCCTTCTCGAACAATACACTACATCTCCAGATAAATAACCATACACTGCACCTCAT	1800
QY	1801	ACTTGCCCGAGGACCCATTAAAGTTGGGGTATTTAAAGTATCTGCCAAATACCAAGGATGAGC	1860
DB	1801	ACTTGCCCGAGGACCCATTAAAGTTGGGGTATTTAAAGTATCTGCCAAATACCAAGGATGAGC	1860
QY	1861	AACAACAGTAAATCAAACTACTGCGCGATTTCCCTGCGACTAGCCGAGGAGCAGGGCTT	1920
DB	1861	AACAACAGTAAATCAAACTACTGCGCGATTTCCCTGCGACTAGCCGAGGAGCAGGGCTT	1920
QY	1921	TGAGACTCCTGTTGGGACACAGATTGGTCTGCGACTCGGCCAGGACGGTCTACTCAGCACAC	1980
DB	1921	TGAGACTCCTGTTGGGACACAGATTGGTCTGCGACTCGGCCAGGACGGTCTACTCAGCACAC	1980
QY	1981	ACTGACTGCTTCAGTGTGCTATCAGAAGATGTCCTCTATCAATTGTGAATGATGGAAAC	2040
DB	1981	ACTGACTGCTTCAGTGTGCTATCAGAAGATGTCCTCTATCAATTGTGAATGATGGAAAC	2040
QY	2041	TTTTAAACCTCCCTCCTCCTCCTTTCACCTCTGCACCTAGTTTTTCCCAATGGTTTC	2100
DB	2041	TTTTAAACCTCCCTCCTCCTCCTTTCACCTCTGCACCTAGTTTTTCCCAATGGTTTC	2100
QY	2101	CAGCAAAAGGTGACTTATTAATATATTTAGTGTGTTTCCAGAAATAAAAAA	2151
DB	2101	CAGCAAAAGGTGACTTATTAATATATTTAGTGTGTTTCCAGAAATAAAAAA	2151
RESULT 3			
AAZ93350			
ID	AAZ93350 standard; cDNA; 2151 BP.		
AC	AAZ93350;		
XX	16-AUG-2000 (first entry)		
XX	Sequence encoding F-box protein FBP-1.		
XX	F-box protein; FBP; diagnosis; treatment; screening; agonist;		
XX	antagonist; proliferative disorder; differentiative disorder;		
XX	breast cancer; prostate cancer; ovarian cancer; cancer;		
XX	small cell lung carcinoma; immune disorder; cardiovascular disorder;		
XX	inflammatory disorder; human; ss.		
OS	Homo sapiens.		
XX	Key		
XX	Location/Qualifiers		
FT	70..1779		
FT	/*tag= a		
FT	/product= FBP-1		
XX	WC200012679-A1.		
XX	27-AUG-1999; 99WO-US19560.		
XX	28-AUG-1998; 98US-0098355.		
XX	03-PES-1999; 99US-0118568.		
XX	15-MAR-1999; 99US-0124449.		
XX	(UYNV) UNIV NEW YORK STATE.		
XX	Chiaur DS, Pagano M, Latres E;		
PI			


```

1741 CGTTCCTCCCTTCGAACTACACCTACATCTCCAGATTAATACCATACATGACCTCAT 1800
1801 ACTTGCCTCCAGACCCATTAAAGTTGGGTAATTAACGTAATCTGCAATACAGGATGAGC 1860
1801 ACTTGCCTCCAGACCCATTAAAGTTGGGTAATTAACGTAATCTGCAATACAGGATGAGC 1860
1861 AACACAGTAACAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGCAGGCGCTT 1920
1861 AACACAGTAACAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGCAGGCGCTT 1920
1921 TGAGACTCTCTGTGGACACAGTTGGTCTGAGTCTGGCCGAGGAGCGGTCTACTCAGACA 1980
1921 TGAGACTCTCTGTGGACACAGTTGGTCTGAGTCTGGCCGAGGAGCGGTCTACTCAGACA 1980
1981 ACTGACTGTCTTCTAGTCTGCTATCAGAGATGCTCTTCTATCAATTTGTAATGTGGAAC 2040
1981 ACTGACTGTCTTCTAGTCTGCTATCAGAGATGCTCTTCTATCAATTTGTAATGTGGAAC 2040
2041 TTTTAACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
2041 TTTTAACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
2101 CAGACAAAGTGACTTATAAATATATTTAGTCTTTTTCAGAAAAA 2151
2101 CAGACAAAGTGACTTATAAATATATTTAGTCTTTTTCAGAAAAA 2151

```

RESULT 4

AAZ93710
ID AAZ93710 standard; DNA; 2151 BP.

AC AAZ93710;

XX AAZ93710;

DT 16-AUG-2000 (first entry)

XX F -box protein hBetaTrCp coding sequence.

XX Ubiquitin ligase; SCF, F-box protein; targeted degradation;
KW destabilitation; proteolysis; drug discovery; gene therapy; cancer;
KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
KW human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 70..1779
CDS

FT /*tag= a
FT /product= F-box protein hBetaTrCp

XX WO200022110-A2.

XX 20-APR-2000..

XX 08-OCT-1999; 99WO-US23705.

XX 09-OCT-1998; 98US-0103787.

XX (HARD) HARVARD COLLEGE.

XX Zhou P, Howley P;

XX WPI; 2000-317970/27.

XX P-PSDB; AAY83250.

XX Targeting degradation of polypeptide useful for treating cancer and
XX other proliferative disorders, involves conjugating polypeptide with
XX ubiquitin protein ligase or inhibiting ubiquitination using organic
XX compound

XX Claim 10; Page 171; 185pp; English.

XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin

CC ligases) which can be used for the targetted degradation of a target
CC polypeptide in vivo. Targetted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or
CC increasing the level of a target polypeptide and for creating and
CC expressing a destabilized polypeptide which is subjected to SCF
CC mediated proteolysis. Degrading any desired protein in a cell is
CC useful for preventing or treating diseases caused by the presence of
CC abnormal amount of the specific polypeptides, for drug discovery and
CC for gene therapy. Diseases treated include cancer, by degradation of
CC oncoproteins, Huntington's disease, other proliferative disorders and
CC microbial infections. The method provides a quick and easy
CC alternative to gene knockout technology. The target polypeptide can
CC be degraded at all stages, or a specific stage, of development in the
CC mature animal.
XX
SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 21; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCCTTGGCTGCGGCTTGGACCAAAAGGGGGCGGCCCGCGGAGAGCGGACCCAGTGCC 60
DB 1 TGCCTTGGCTGCGGCTTGGACCAAAAGGGGGCGGCCCGCGGAGAGCGGACCCAGTGCC 60

QY 61 TCGGCGATTATGACCCCGCGGAGGCGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGAAT 120
DB 61 TCGGCGATTATGACCCCGCGGAGGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAAT 120

QY 121 TCCTCAGAGAGAGAGACTGTATATATGGGGAACCCCTTAGGAAGATAATACAGAGAAG 180
DB 121 TCCTCAGAGAGAGAGACTGTATATATGGGGAACCCCTTAGGAAGATAATACAGAGAAG 180

QY 181 AATTCACTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
DB 181 AATTCACTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240

QY 241 TGTTACAGAGACTGTATGAAGACTGTGAATTTGTGGCCAAACAAACATCCCAAT 300
DB 241 TGTTACAGAGACTGTATGAAGACTGTGAATTTGTGGCCAAACAAACATCCCAAT 300

QY 301 GGCATTCCAGTATGATTGTGCCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAGGAA 360
DB 301 GGCATTCCAGTATGATTGTGCCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAGGAA 360

QY 361 AAGGAACCTGTGTCAAATACCTTTGAGCGAGTGGTCAAGTCAAGTCAAGTGAATTTGTG 420
DB 361 AAGGAACCTGTGTCAAATACCTTTGAGCGAGTGGTCAAGTCAAGTCAAGTGAATTTGTG 420

QY 421 GAACATCTTATATCCCAAAATGTCTCATTACCAACATGGGCAATAACTCGTATCTTAA 480
DB 421 GAACATCTTATATCCCAAAATGTCTCATTACCAACATGGGCAATAACTCGTATCTTAA 480

QY 481 CCTATGTGCAGAGAGATTTCATAAGTCTGCCAGCTCGGGGATTTGGATCATTCGCT 540
DB 481 CCTATGTGCAGAGAGATTTCATAAGTCTGCCAGCTCGGGGATTTGGATCATTCGCT 540

QY 541 GAGAACATTTCTGTACACCTGGATGCAAAATCACTATGTGCTGCTGAACCTTGTGCAAG 600
DB 541 GAGAACATTTCTGTACACCTGGATGCAAAATCACTATGTGCTGCTGAACCTTGTGCAAG 600

QY 601 GAATGGTACCGAGTACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAAATGGTC 660
DB 601 GAATGGTACCGAGTACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAAATGGTC 660

QY 661 AGGACAGATTCTCTGTGGAGAGGCGCTTGGCAGAACCGAGAGGATGGGACAGTATTATTC 720
DB 661 AGGACAGATTCTCTGTGGAGAGGCGCTTGGCAGAACCGAGAGGATGGGACAGTATTATTC 720

QY 721 AAAAACAACCTCTGTGACGGGAATGCTCTCCCACTCTTTTATAGAGCACTTATGCT 780
DB 721 AAAAACAACCTCTGTGACGGGAATGCTCTCCCACTCTTTTATAGAGCACTTATGCT 780


```

1b 721 AAAAAAACCCTCTGACGGGAAGCTCCTCCCAACTCTTTTATAGAGCACTTTATCCT 780
1b 781 AAAATTTATACAGACATTTGAGACAAATGAATCTAATTTGGAGATGTGGAACATAGTTTA 840
1b 781 AAAATTTATACAGACATTTGAGACAAATGAATCTAATTTGGAGATGTGGAACATAGTTTA 840
1b 841 CAGAGAAATTCATGCGGAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGAT 900
1b 841 CAGAGAAATTCATGCGGAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGAT 900
1b 901 CAGAAAAATAGTAAGCGGCTTCGACACAAACAAATCAAGATCTGGGATAAAAAACATTC 960
1b 901 CAGAAAAATAGTAAGCGGCTTCGACACAAACAAATCAAGATCTGGGATAAAAAACATTC 960
1b 961 GAATGCAAGCAATTCATCAGAGGCAATACAGGTTCAAGTCTCTCTCCAGTATGATGAG 1020
1b 961 GAATGCAAGCAATTCATCAGAGGCAATACAGGTTCAAGTCTCTCTCCAGTATGATGAG 1020
1b 1021 AGAGTGATCATACAGGATCATCGGATTCACGGTCAAGGTGCGGATGTAATACAGGT 1080
1b 1021 AGAGTGATCATACAGGATCATCGGATTCACGGTCAAGGTGCGGATGTAATACAGGT 1080
1b 1081 GAAATGCTAAACACAGTTGATTCACCAATGTAAGCAAGTCTGCACTTGCCTTTCAATAAT 1140
1b 1081 GAAATGCTAAACACAGTTGATTCACCAATGTAAGCAAGTCTGCACTTGCCTTTCAATAAT 1140
1b 1141 GGCATGATGTCACCTGCTCCAAAGATCGTTCCATTCGATGATGCGGATGCGCTCCCA 1200
1b 1141 GGCATGATGTCACCTGCTCCAAAGATCGTTCCATTCGATGATGCGGATGCGCTCCCA 1200
1b 1201 ACTGACATTAACCTCCGAGGAGTGTGTCGGAACACGAGTGTCTGCAATGTTGTAGAC 1260
1b 1201 ACTGACATTAACCTCCGAGGAGTGTGTCGGAACACGAGTGTCTGCAATGTTGTAGAC 1260
1b 1261 TTTGATGCAAGTACATGTTTCTGCATCTGGGATGAACTATAAAGGTATGGAACACA 1320
1b 1261 TTTGATGCAAGTACATGTTTCTGCATCTGGGATGAACTATAAAGGTATGGAACACA 1320
1b 1321 AGTACTGTGAATTTGTAAGGACCTTAAATGAGCAAAACGAGGCAATGCGTGTTCAG 1380
1b 1321 AGTACTGTGAATTTGTAAGGACCTTAAATGAGCAAAACGAGGCAATGCGTGTTCAG 1380
1b 1381 TACAGGAGCAGCTGTGATGAGTGGCTCATCTGACACACTATCAGATTATGGACATA 1440
1b 1381 TACAGGAGCAGCTGTGATGAGTGGCTCATCTGACACACTATCAGATTATGGACATA 1440
1b 1441 GAATGTGTCATGTTTACAGTGTGTAGAGGCCATGAGGAATGCGTGTGATTCGA 1500
1b 1441 GAATGTGTCATGTTTACAGTGTGTAGAGGCCATGAGGAATGCGTGTGATTCGA 1500
1b 1501 TTTGATACAAAGAGTATGATGAGTGGGCTATGATGGAATAATTAAGTGTGGATCTT 1560
1b 1501 TTTGATACAAAGAGTATGATGAGTGGGCTATGATGGAATAATTAAGTGTGGATCTT 1560
1b 1561 GTGCTGCTTTGGACCCCGCTCTCTGAGGACACTCTGTCACGACCTTGTGGAG 1620
1b 1561 GTGCTGCTTTGGACCCCGCTCTCTGAGGACACTCTGTCACGACCTTGTGGAG 1620
1b 1621 CATTCGGAAGATTTTCGACTACATGTTGATGATTCAGATTCAGATTCAGATTCAGAT 1680
1b 1621 CATTCGGAAGATTTTCGACTACATGTTGATGATTCAGATTCAGATTCAGATTCAGAT 1680
1b 1681 GATGACACAAATCTCATCTGGGACTTCTTAATGATCCAGTGTGCGCAAGCTGAACCCCT 1740
1b 1681 GATGACACAAATCTCATCTGGGACTTCTTAATGATCCAGTGTGCGCAAGCTGAACCCCT 1740
1b 1741 CGTTCCCTCTTCGAAACATACACCTACATCTCCAGATAAATCAATACATGACCTCAT 1800
1b 1741 CGTTCCCTCTTCGAAACATACACCTACATCTCCAGATAAATCAATACATGACCTCAT 1800
1b 1801 ACTTGGCCAGGACCCATTAAGTGTGGTATTAACGATCTGCGCAATACAGGATGAGC 1860
1b 1801 ACTTGGCCAGGACCCATTAAGTGTGGTATTAACGATCTGCGCAATACAGGATGAGC 1860

```

```

RESULT 5
AAC84610
ID AAC84610 standard; DNA; 2151 BP.
XX
AC AAC84610;
XX
DT 02-APR-2001 (first entry)
XX
DE Human ZF11 protein encoding DNA.
XX
KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
Bad; Bcl-2; tumour; cytosstatic; ds.
XX
OS Homo sapiens.
XX
PN WO200075184-A1.
XX
PD 14-DEC-2000.
XX
PF 05-JUN-2000; 2000WO-US15449.
XX
PR 04-JUN-1999; 99US-0137494.
XX
PA (UYVA ) UNIV YALE.
XX
PI Zhang H, Tsvetkov LM, Kondo T;
XX
DR WPI: 2001-061703/07.
XX
P-PSDB; AAB48298.
XX
PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
involves altering levels of proteins such as S-phase kinase associated
proteins 1, 2 and cullin/CDC53 proteins -
XX
PS Examples; Page 129-130; 162pp; English.
XX
CC The invention relates to methods of altering the polypeptide levels in a
cell, using proteins selected from S-phase kinase associated proteins 1
and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
cullin/CDC53 family of proteins). The method is useful for altering the
level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
detecting tumours, and in monitoring tumor treatment in a mammal. Agents
that modulate interactions between SKP and target proteins are useful for
treating tumours.
XX
SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

```

```

Query Match 100.0%; Score 2151; DB 22; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;

```

Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
2y	1	TGGTTGGCTGCGGCTGGCAACAAAGGGGGCGGCCCGGGAGACGGACCCAGTGCC	60
Db	1	TGGTTGGCTGCGGCTGGCACAAAGGGGGCGGCCCGGGAGACGGACCCAGTGCC	60
2y	61	TGGCGAATTATGGAACCGGCGAGCGGGTGTGCAGAGAAGGCATCAAGTTTATGAAT	120
Db	61	TGGCGAATTATGGAACCGGCGAGCGGGTGTGCAGAGAAGGCATCAAGTTTATGAAT	120
2y	121	TCCTCAGAGAGAAGACTGTAAATAATGGCGAACCCCTAGGAAGATAATACGAGAAAG	180
Db	121	TCCTCAGAGAGAAGACTGTAAATAATGGCGAACCCCTAGGAAGATAATACGAGAAAG	180
2y	181	AATTCACTTAGACAGACATACAAACAGCTGTGCCAGACTCTGCTTAAACAAGAAACAGTA	240
Db	181	AATTCACTTAGACAGACATACAAACAGCTGTGCCAGACTCTGCTTAAACAAGAAACAGTA	240
2y	241	TGTTTAGCAAGCACTGCTATGAAGACTTGAGAAATTTGTGGCCAAACAACCTTCCCAAT	300
Db	241	TGTTTAGCAAGCACTGCTATGAAGACTTGAGAAATTTGTGGCCAAACAACCTTCCCAAT	300
2y	301	GGCACTTCCAGTATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATCAAAAGGAA	360
Db	301	GGCACTTCCAGTATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATCAAAAGGAA	360
2y	361	AAGGAACCTGTGTCAAAATCTTTGACAGCTGTGCAGATCAGATCAAGTGAATTTGTG	420
Db	361	AAGGAACCTGTGTGTCAAAATCTTTGACAGCTGTGCAGATCAGATCAAGTGAATTTGTG	420
2y	421	GAACATCTTATATCCAAATGTGCATTACCAACATGGGCACATAAATCGTATCTTAAA	480
Db	421	GAACATCTTATATCCAAATGTGCATTACCAACATGGGCACATAAATCGTATCTTAAA	480
2y	481	CCTATGTTCCAGAGAAATTCAATACTGTCTGCGAGCTCGGGATTTGGATCAATGCGT	540
Db	481	CCTATGTTCCAGAGAAATTCAATACTGTCTGCGAGCTCGGGATTTGGATCAATGCGT	540
2y	541	GAGAAACATTTCTGTACTCTGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTGAAG	600
Db	541	GAGAAACATTTCTGTACTCTGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTGAAG	600
2y	601	GAATGTGACCGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCAGAGAAATGTC	660
Db	601	GAATGTGACCGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCAGAGAAATGTC	660
2y	661	AGGACAGATTCTCTGTGAGAGGCTGCGCAGACGAAGAGAGATGGGACAGTATTATTTC	720
Db	661	AGGACAGATTCTCTGTGAGAGGCTGCGCAGACGAAGAGAGATGGGACAGTATTATTTC	720
2y	721	AAAAACAAACCTCTGTGACGGGAATGCTCCTCCCAACTCTTTTATAGACACTTTATCCT	780
Db	721	AAAAACAAACCTCTGTGACGGGAATGCTCCTCCCAACTCTTTTATAGACACTTTATCCT	780
2y	781	AAAAATTATACAGACATTGACACAATAGAACTCTAAATGGAGATGTGGAGACATAGTTTA	840
Db	781	AAAAATTATACAGACATTGACACAATAGAACTCTAAATGGAGATGTGGAGACATAGTTTA	840
2y	841	CAGAGAAATTCACGCCGAAGTGAACAAAGCAAAAGGAGTTTACTGTTTACAGTATGATGAT	900
Db	841	CAGAGAAATTCACGCCGAAGTGAACAAAGCAAAAGGAGTTTACTGTTTACAGTATGATGAT	900
2y	901	CAGAAANTAGTAAGCGGCTTTGAGACAAACAAATCAAGATCTGGGATAAAAACATTTG	960
Db	901	CAGAAANTAGTAAGCGGCTTTGAGACAAACAAATCAAGATCTGGGATAAAAACATTTG	960
2y	961	GAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTTCAGTCTCTCTCAGTATGATGAG	1020
Db	961	GAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTTCAGTCTCTCTCAGTATGATGAG	1020
2y	1021	AGATGATCATTAACAGGATTCATCGAATTCACGGTCCAGAGTGTGGATGTAATACAGGT	1080
Db	1021	AGATGATCATTAACAGGATTCATCGAATTCACGGTCCAGAGTGTGGATGTAATACAGGT	1080

Qy	1081	GAATGCTAAACAGCTTGAATTCACCAATGTGAAGACAGTTCTGCACCTCGGTTTCAATAAT	1114
Db	1081	GAATGCTAAACAGCTTGAATTCACCAATGTGAAGACAGTTCTGCACCTCGGTTTCAATAAT	1114
Qy	1141	GGCATGATGGTGACCTGCTCCAAAGATTCGTTCCATTGCTGATGGGATATGGCCTCCCCCA	1200
Db	1141	GGCATGATGGTGACCTGCTCCAAAGATTCGTTCCATTGCTGATGGGATATGGCCTCCCCCA	1200
Qy	1201	ACTGACATTACCTCCGGAGGGTGTGTGGACACCGAGCTGTCTCAATGTTGTAGAC	1260
Db	1201	ACTGACATTACCTCCGGAGGGTGTGTGGACACCGAGCTGTCTCAATGTTGTAGAC	1260
Qy	1261	TTTGATGCAAGTACATTGTTTCTGCATTCTGGGGATAGAACTATAAAGGATATGGAACACA	1320
Db	1261	TTTGATGCAAGTACATTGTTTCTGCATTCTGGGGATAGAACTATAAAGGATATGGAACACA	1320
Qy	1321	AGTACTTGTGAATTTGTAAGGACCTTAAATGAGACACAAACGAGGCAATTGGCTGTTTCGAG	1380
Db	1321	AGTACTTGTGAATTTGTAAGGACCTTAAATGAGACACAAACGAGGCAATTGGCTGTTTCGAG	1380
Qy	1381	TACAGGACAGCGCTGGTAGTGAGTGGCTCATCTGACCAACACTATCAGATTATGGGACATA	1440
Db	1381	TACAGGACAGCGCTGGTAGTGAGTGGCTCATCTGACCAACACTATCAGATTATGGGACATA	1440
Qy	1441	GAATGGTGACATGTTTACAGATGTTTAGAGGCCATGAGGAATTTGGTGCCTGTATTTCGA	1500
Db	1441	GAATGGTGACATGTTTACAGATGTTTAGAGGCCATGAGGAATTTGGTGCCTGTATTTCGA	1500
Qy	1501	TTTGATAACAAGAGGATAGTCAGTGGGGCCTATGATGGAAAAATTAAGTGTGGGATCTT	1560
Db	1501	TTTGATAACAAGAGGATAGTCAGTGGGGCCTATGATGGAAAAATTAAGTGTGGGATCTT	1560
Qy	1561	GTGGTGCTTTGGACCCCGTGCTCTGACGAGCACTCTGTCTACGGACCCCTGTGGAG	1620
Db	1561	GTGGTGCTTTGGACCCCGTGCTCTGACGAGCACTCTGTCTACGGACCCCTGTGGAG	1620
Qy	1621	CATTCGGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCACTAGTTTCACAT	1680
Db	1621	CATTCGGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCACTAGTTTCACAT	1680
Qy	1681	GATGACAAATTCCTCATCTCGGACTTCCCTAAATGATCCAGCTGCCCAGCTGAACCCCC	1740
Db	1681	GATGACAAATTCCTCATCTCGGACTTCCCTAAATGATCCAGCTGCCCAGCTGAACCCCC	1740
Qy	1741	CGTTCGCCCTTCTCGAACATACACTACATCTCCAGATAAATAACCATACACTGACCTCAT	1800
Db	1741	CGTTCGCCCTTCTCGAACATACACTACATCTCCAGATAAATAACCATACACTGACCTCAT	1800
Qy	1801	ACTTCCGCCAGAGCCATTAAAGTTGGGTATTTTACGTTCTGGCAATACACGATGAGC	1860
Db	1801	ACTTCCGCCAGAGCCATTAAAGTTGGGTATTTTACGTTCTGGCAATACACGATGAGC	1860
Qy	1861	AAACAACAGTAAACAATCAAACTA	1920
Db	1861	AAACAACAGTAAACAATCAAACTA	1920
Qy	1921	TCAGACTCCTGTTGGGACACAGTTGTCTGCTGACGTCGGCCAGGACGGTCTACTCAGACACA	1980
Db	1921	TCAGACTCCTGTTGGGACACAGTTGTCTGCTGACGTCGGCCAGGACGGTCTACTCAGACACA	1980
Qy	1981	ACTGACTGCTTCACTGAGTGTCTATCAGAGATGCTTCTTATCAATTTGTAATTTGGAAC	2040
Db	1981	ACTGACTGCTTCACTGAGTGTCTATCAGAGATGCTTCTTATCAATTTGTAATTTGGAAC	2040
Qy	2041	TTTTTAAACCTCCCTCCTCTCCCTTTTCACTCTGACCTAGTTTTTTTCCCATTTGGTTC	2100
Db	2041	TTTTTAAACCTCCCTCCTCTCCCTTTTCACTCTGACCTAGTTTTTTTCCCATTTGGTTC	2100
Qy	2101	CAGACAAAGGTGACTTATAAATATATTTTAGTGTTTTGGCAGAAAAA	2151
Db	2101	CAGACAAAGGTGACTTATAAATATATTTTAGTGTTTTGGCAGAAAAA	2151

Db 1261 TTTGATGACAAAGTACATCTGTTCTGCACTGCGGATAGAACTATAAGGTATGACACA 1320
 Qy 1321 AGTACTTGTGAATTTCTAGGACCTTAATGACACAAACGAGCAATGCTGTTTCAG 1380
 Db 1321 AGTACTTGTGAATTTCTAGGACCTTAATGACACAAACGAGCAATGCTGTTTCAG 1380
 Qy 1381 TACAGGACAGGCTGTAGTGTGCTCATCTGCAACACATATCAGATTATGGACATA 1440
 Db 1381 TACAGGACAGGCTGTAGTGTGCTCATCTGCAACACATATCAGATTATGGACATA 1440
 Qy 1441 GATGTGGTGCATTTTACGAGTGTAGAGCCATGAGGAATGGTCGTTGATTCGA 1500
 Db 1441 GATGTGGTGCATTTTACGAGTGTAGAGCCATGAGGAATGGTCGTTGATTCGA 1500
 Qy 1501 TTTGATAACAAGAGGATAGTCACTGAGTGGGCTTATGATGGAATAAAGTGTGGATCTT 1560
 Db 1501 TTTGATAACAAGAGGATAGTCACTGAGTGGGCTTATGATGGAATAAAGTGTGGATCTT 1560
 Qy 1561 GTGGCTGTTTGGACCCCGTCTCTGAGGACACTCTGCTACGACCCCTTGTGGAG 1620
 Db 1561 GTGGCTGTTTGGACCCCGTCTCTGAGGACACTCTGCTACGACCCCTTGTGGAG 1620
 Qy 1621 CATTCGGGAAGATTTTTCGACTACAGTTTGATGAATCCAGATTGTCAGTAGTTCACAT 1680
 Db 1621 CATTCGGGAAGATTTTTCGACTACAGTTTGATGAATCCAGATTGTCAGTAGTTCACAT 1680
 Qy 1681 GATGACACATCTCATCTGAGCACTCTTAATGATCCAGTGTCCAGCTGAACCCGCC 1740
 Db 1681 GATGACACATCTCATCTGAGCACTCTTAATGATCCAGTGTCCAGCTGAACCCGCC 1740
 Qy 1741 CGTTCCTCTTTCGAAATACACCTACATCTCCAGATAAATACATACAGTCACTCAT 1800
 Db 1741 CGTTCCTCTTTCGAAATACACCTACATCTCCAGATAAATACATACAGTCACTCAT 1800
 Qy 1801 ACTTGCCAGGACCCATTAAGTGGGTATTAACGTATCTGCAATACAGGATGAGC 1860
 Db 1801 ACTTGCCAGGACCCATTAAGTGGGTATTAACGTATCTGCAATACAGGATGAGC 1860
 Qy 1861 AACACAGTAACAATCAAACTACTGCCAGTTTCCCTGACTAGCCGAGGACAGGCTT 1920
 Db 1861 AACACAGTAACAATCAAACTACTGCCAGTTTCCCTGACTAGCCGAGGACAGGCTT 1920
 Qy 1921 TGAGACTCTGTTGGACACATTTGGTGTGCACTGCGCCAGGACGGTCTACTCAGACA 1980
 Db 1921 TGAGACTCTGTTGGACACATTTGGTGTGCACTGCGCCAGGACGGTCTACTCAGACA 1980
 Qy 1981 ACTGACTGCTTCAGTCTCTATCAGAAGATGCTCTCTATCAATTGTGAATTTGAAC 2040
 Db 1981 ACTGACTGCTTCAGTCTCTATCAGAAGATGCTCTCTATCAATTGTGAATTTGAAC 2040
 Qy 2041 TTTTAACTCCCT 2100
 Db 2041 TTTTAACTCCCT 2100
 Qy 2101 CAGACAAAGTGACTATATAATATTTAGTGTGTTTTCAGAAAAA 2151
 Db 2101 CAGACAAAGTGACTATATAATATTTAGTGTGTTTTCAGAAAAA 2151

RESULT 7

ABSS1009

ID ABSS1009 standard; cDNA; 2151 BP.

XX 4C

XX ABSS1009;

XX 21-OCT-2002 (first entry)

XX Human cDNA encoding bait protein beta-TrCP1.

XX Human; ss; gene; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;

XX non-insulin diabetes mellitus; obesity; selected interacting domain;

XX SID; protein-protein interaction map; PIM; anorectic; metabolic disorder.

XX

OS Homo sapiens.
 XX WO200253726-A2.
 PN 11-JUL-2002.
 PD 28-DEC-2001; 2001WO-EPI5423.
 PF 02-JAN-2001; 2001US-259377P.
 PR (HYBR-) HYBRIGENICS.
 XX (CNRS) CENT NAT RECH SCI.
 PA Legrain P, Marullo S, Jockers R;
 PI WPI: 2002-583612/62.
 XX P-PSDB; ABG69473.

Novel complex of protein-protein interactions in adipocyte cells for identifying compounds that modulate the protein-protein interactions and useful for treating obesity and metabolic disorders

Claim 2; Page -; 125pp; English.

The invention relates to a complex of protein-protein interactions (forming a protein-protein interaction map, PIM) in adipocyte cells as defined in the specification, or polynucleotides in adipocytes encoding for the polypeptides. Also included are a recombinant cell expressing the interacting polypeptides and a method of selecting a modulating compound in adipocyte cells, by cultivating a recombinant host cell on a selective medium containing a modulating compound and a reporter gene the expression of which is toxic for the recombinant host cell which is transformed with two vectors, where the first vector comprises a polynucleotide encoding a first hybrid polypeptide and DNA binding domain and the second vector comprising a polynucleotide encoding a second hybrid polypeptide and an activating domain that activates the toxic reporter gene, when the first and second hybrid polypeptides interact and selecting the modulating compound which inhibits the growth of the recombinant host cell (i.e. using the yeast two-hybrid system). The complexes are useful for identifying compounds that modulate the protein-protein interactions and useful for treating obesity and metabolic disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The compound isolated by the method is useful for treating and preventing obesity or metabolic diseases. The interactions between the proteins of the complex further define a set of selected interacting domains, SID. The present sequence encodes a member of the protein complex of the invention, used as the bait protein in the yeast two-hybrid assay.

Note: The present sequence was not displayed in the specification but was obtained from its Genbank entry by the indexer.

Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 24; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTGGCTGCGGCTGCGACCAAGGGGCGGCCCGCGGAGAGCGACCCAGTGCC 60
 Db 1 TGCCTTGGCTGCGGCTGCGACCAAGGGGCGGCCCGCGGAGAGCGACCCAGTGCC 60
 Qy 61 TCGCGCATTTATGACCCCGCGGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAAT 120
 Db 61 TCGCGCATTTATGACCCCGCGGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAAT 120
 Qy 121 TCCTCAGAGAGAGAGACTGTATAATGCGCAACCCCTTAGGAAGATAATACAGAGAAG 180
 Db 121 TCCTCAGAGAGAGAGACTGTATAATGCGCAACCCCTTAGGAAGATAATACAGAGAAG 180
 Qy 181 AATTCACCTTAGACAGACATACACACTGTGCCAGACTGTGCTTAAACCAAGAACAGTA 240
 Db 181 AATTCACCTTAGACAGACATACACACTGTGCCAGACTGTGCTTAAACCAAGAACAGTA 240

Db 491 AGAGAGATTTCATAACTGCTCTGCGAGCTCGGGGATTGGATCATATTGCTGAGAACATTC 550
2y 551 TGTCTATACCTGGATGCCAAATCACTATATGTGCTGCTGAACCTTGTGCAAGGAATGGTACC 610
Db 551 TGTCTATACCTGGATGCCAAATCACTATATGTGCTGCTGAACCTTGTGCAAGGAATGGTACC 610
2y 611 GAGTGACCTCTGATGCGATGCTGTGGAGAGACTTATCGAGAGATGGTCAGACAGATT 670
Db 611 GAGTGACCTCTGATGCGATGCTGTGGAGAGACTTATCGAGAGATGGTCAGACAGATT 670
2y 671 CTCTGTGGAGAGCCCTGGCAGAACGAAGAGATGGGACAGTATTATTCAAAAACAAC 730
Db 671 CTCTGTGGAGAGCCCTGGCAGAACGAAGAGATGGGACAGTATTATTCAAAAACAAC 730
2y 731 CTCTGAGCGGATGCTCTCCCACTCTTTTATAGAGCACTTATCCTAAATATTATC 790
Db 731 CTCTGAGCGGATGCTCTCCCACTCTTTTATAGAGCACTTATCCTAAATATTATC 790
2y 791 AAGACATTGAGACAATAGAAATCTAATTGGAGATGTGGAAGACATAGTTTACAGAGAATTC 850
Db 791 AAGACATTGAGACAATAGAAATCTAATTGGAGATGTGGAAGACATAGTTTACAGAGAATTC 850
2y 851 ACTGCCGAGTGAACAAGCAAGAGTTTACTGTTTACAGTATGATGATCAGAGATAG 910
Db 851 ACTGCCGAGTGAACAAGCAAGAGTTTACTGTTTACAGTATGATGATCAGAGATAG 910
2y 911 TAAGCGGCCCTTCGAGACAACAATCAAGATCTGGGATAAAAACACATTTGGAATGCTAA 970
Db 911 TAAGCGGCCCTTCGAGACAACAATCAAGATCTGGGATAAAAACACATTTGGAATGCTAA 970
2y 971 GAATTTCTCAGAGCCATAGCTTCACTCTCTGCTCCAGTATGATGAGAGATGATCA 1030
Db 971 GAATTTCTCAGAGCCATAGCTTCACTCTCTGCTCCAGTATGATGAGAGATGATCA 1030
2y 1031 TAACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTAATACAGGTGAATGCTAA 1090
Db 1031 TAACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTAATACAGGTGAATGCTAA 1090
2y 1091 ACAGTTGATTCACCATTTGAGAGAGTTCTGCACTTGGCTTCAATTAATGGCATGTTG 1150
Db 1091 ACAGTTGATTCACCATTTGAGAGAGTTCTGCACTTGGCTTCAATTAATGGCATGTTG 1150
2y 1151 TGACCTCTCCAAAGATCGTTCCATTCGTATGGGATATGGCTCCCACTGACATTA 1210
Db 1151 TGACCTCTCCAAAGATCGTTCCATTCGTATGGGATATGGCTCCCACTGACATTA 1210
2y 1211 CCCTCGGAGGTGCTGGTCCGACACAGAGCTGCTCAATGTTGATGATGATGACA 1270
Db 1211 CCCTCGGAGGTGCTGGTCCGACACAGAGCTGCTCAATGTTGATGATGATGACA 1270
2y 1271 AGTACATTTGTTCTGCTGCTGGGATAGAACTATAAGGTATGGAACAAAGTCTTGTG 1330
Db 1271 AGTACATTTGTTCTGCTGCTGGGATAGAACTATAAGGTATGGAACAAAGTCTTGTG 1330
2y 1331 AATTGTAAGGACCTTAATAGGACAAACAGAGGATTCCTGTTGCGATACAGGACA 1390
Db 1331 AATTGTAAGGACCTTAATAGGACAAACAGAGGATTCCTGTTGCGATACAGGACA 1390
2y 1391 GGCTGGTAGTGAGTGCATCTGACAAACACATATAGATTTAGGACATAGAAATGGTG 1450
Db 1391 GGCTGGTAGTGAGTGCATCTGACAAACACATATAGATTTAGGACATAGAAATGGTG 1450
2y 1451 CATGTTTACAGTGTATAGAGCCATGAGGAATGGTGGTGTATTCGATTTGATACAA 1510
Db 1451 CATGTTTACAGTGTATAGAGCCATGAGGAATGGTGGTGTATTCGATTTGATACAA 1510
2y 1511 AGAGATAGTCAGTGGGCTTATGATGGAATAATTAAGTGTGGATTTTGTGGCTGCTT 1570
Db 1511 AGAGATAGTCAGTGGGCTTATGATGGAATAATTAAGTGTGGATTTTGTGGCTGCTT 1570
2y 1571 TGGACCCCTGCTCTGCGAGGACACTCTGCTACGGACCTCTGCGGATTTCCGGA 1630
Db 1571 TGGACCCCTGCTCTGCGAGGACACTCTGCTACGGACCTCTGCGGATTTCCGGA 1630

QY 1631 GAGTTTTGCACTACAGTTTGAATATCCAGATTGTCAAGATTGTCAATGATGACACAA 1690
Db 1631 GAGTTTTGCACTACAGTTTGAATATCCAGATTGTCAAGATTGTCAATGATGACACAA 1690
QY 1691 TCCTCATCTGGGACTTCCATAATGATCCAGCTGCCCAAGCTGAACCCCGTTCCCTTT 1750
Db 1691 TCCTCATCTGGGACTTCCATAATGATCCAGCTGCCCAAGCTGAACCCCGTTCCCTTT 1750
QY 1751 CTGGAACATACACCTACATCTCCAGATAAATAACCATACATGACCTCATACTTGGCCAG 1810
Db 1751 CTGGAACATACACCTACATCTCCAGATAAATAACCATACATGACCTCATACTTGGCCAG 1810
QY 1811 GACCCATTAAGTTGGGTATTTAAAGTATCTGCCAATACCAGGATGAGCAACAAGTA 1870
Db 1811 GACCCATTAAGTTGGGTATTTAAAGTATCTGCCAATACCAGGATGAGCAACAAGTA 1870
QY 1871 ACAATCAAACTACGCCCAGTTTCCCTGGACTAGCCGAGGAGCGGCTTTGAGACTCCT 1930
Db 1871 ACAATCAAACTACGCCCAGTTTCCCTGGACTAGCCGAGGAGCGGCTTTGAGACTCCT 1930
QY 1931 GTTGGGACACAGTTGCTGCGAGTGGCCAGGACGGTCTACTCAGCAACAACCTGACTGCT 1990
Db 1931 GTTGGGACACAGTTGCTGCGAGTGGCCAGGACGGTCTACTCAGCAACAACCTGACTGCT 1990
QY 1991 TCAGTGTCTATCAGAGATGCTTTCTATCAATTTGTAATGATGGAACTTTTAAACCT 2050
Db 1991 TCAGTGTCTATCAGAGATGCTTTCTATCAATTTGTAATGATGGAACTTTTAAACCT 2050
QY 2051 CCCCTCTCTCTCCCTTTTCCACCTGCGACTAGTTTTCCTATTTTCCATTTGTTCCAGACAAAGG 2110
Db 2051 CCCCTCTCTCTCCCTTTTCCACCTGCGACTAGTTTTCCTATTTTCCATTTGTTCCAGACAAAGG 2110
QY 2111 TGACTTATAAATATATTAGTGTGTTT 2136
Db 2111 TGACTTATAAATATATTAGTGTGTTT 2136

RESULT 10
AAH89966
ID AAH89966 standard; cDNA; 3622 BP.
XX AAH89966;
XX AC
XX DT 01-OCT-2001 (first entry)
XX Human bone marrow cDNA, SEQ ID NO: 97.
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
OS Homo sapiens.
XX
XX WO200153453-A2.
XX
XX 26-JUL-2001.
XX
XX 23-DEC-2000; 2000WO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.

1 Ford JE, Boyle BU, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
2 Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
3 Zhao QA, Zhou P, Drmanac RT;

4 WPI; 2001-488707/53.
5 P-PSDB; AAM00847.

6 Novel bone-marrow-expressed polynucleotides and polypeptides, useful
7 for treating e.g. cancer and immune deficiency disorders -

8 Claim 1; Page 274-275; 648pp; English.

9 The present sequence is one of 251 novel human polynucleotides
10 expressed in the bone marrow. The polynucleotide and the
11 polypeptide encoded by it are useful in the treatment of various
12 immune deficiencies and disorders. The deficiencies and disorders may
13 be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
14 infection, or may result from an autoimmune disorder, a coagulation
15 disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
16 suppression of an inflammatory response or treatment of a nervous
17 system disorder such as Alzheimer's disease. Detection of the presence
18 or increased expression of the polynucleotide or the protein it
19 encodes is useful for the diagnosis and/or prognosis of one
20 or more types of cancer. The polynucleotide and polypeptide can be
21 used as nutritional sources or supplements and in the screening of
22 chemical compounds as potential drugs.

23 Sequence 3622 BP; 961 A; 839 C; 874 G; 948 T; 0 other;

24 Query Match 93.5%; Score 2011.4; DB 22; Length 3622;
25 Best Local Similarity 99.7%; Pred. No. 0;
26 Matches 2026; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

27 112 TTTATGAATTCCTCAGAGAGAACTGTAAATGCGAACCCCTAGAGATAATA 171
28 261 TTCCAGAAATTCCTCAGAGAGAGAACTGTAAATGCGAACCCCTAGAGATAATA 320
29 172 CAGAGAGAAATTCCTCAGAGAGAGAACTGTAAATGCGAACCCCTAGAGATAATA 231
30 321 CAGAGAGAAATTCCTCAGAGAGAGAACTGTAAATGCGAACCCCTAGAGATAATA 380
31 232 GAAACAGATGTTTACGAGAGAACTGTAAATGCGAACCCCTAGAGATAATA 291
32 381 GAAACAGATGTTTACGAGAGAACTGTAAATGCGAACCCCTAGAGATAATA 440
33 292 CTTGCCAATGGCACTTCCAGTATGATTTGCCAAGCAACGAACTCTCAGCAAGCTAT 351
34 441 CTTGCCAATGGCACTTCCAGTATGATTTGCCAAGCAACGAACTCTCAGCAAGCTAT 500
35 352 GAAAGGAAAGGAACTGTGTCAAAATCTTTCAGAGAGTGGTTCAGAGTCAAGTGC 411
36 501 GAAAGGAAAGGAACTGTGTCAAAATCTTTCAGAGAGTGGTTCAGAGTCAAGTGC 560
37 412 GAATTTGTGGAACATCTTATATCCAAATGTGTCAATACCAACATGGGCACATAAATCG 471
38 561 GAATTTGTGGAACATCTTATATCCAAATGTGTCAATACCAACATGGGCACATAAATCG 620
39 472 TATCTTAACCTATGTCAGAGAGATTTCAATCTCTGCCAGTCTGCCAGTTCGGGATGGAT 531
40 621 TATCTTAACCTATGTCAGAGAGATTTCAATCTCTGCCAGTCTGCCAGTTCGGGATGGAT 680
41 532 CATATCGCTGAGAACTTCTGTCAATCTGGATGCAATCACTATGTGTCTGTGAACCTT 591
42 681 CATATCGCTGAGAACTTCTGTCAATCTGGATGCAATCACTATGTGTCTGTGAACCTT 740
43 592 GTGTGCAAGGATGTTACCGAGTACCTCTGATGGCATGCTGTGGAAAGAGCTTATCGAG 651
44 741 GTGTGCAAGGATGTTACCGAGTACCTCTGATGGCATGCTGTGGAAAGAGCTTATCGAG 800
45 652 AGAATGGTCAGGACAGATTTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAG 711
46 801 AGAATGGTCAGGACAGATTTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAG 860

QY 712 TATTTATTCAAAAACAAACCTCTCTGACGGGAATGCTCTCTCCCAACTCTTTTATAGAGCA 771
DB 861 TATTTATTCAAAAACAAACCTCTCTGACGGGAATGCTCTCTCCCAACTCTTTTATAGAGCA 920
QY 772 CTTTATCTCTAAAATTTACAGACATTTGAGACATTAAGATCTAAATGGAGATGTCGAAGA 831
DB 921 CTTTATCTCTAAAATTTACAGACATTTGAGACATTAAGATCTAAATGGAGATGTCGAAGA 980
QY 832 CATAGTTTACAGAGAAATTCATCTCCGAAGTGAAACAAAGCAAGAGGATTTACTGTTTACAG 891
DB 981 CATAGTTTACAGAGAAATTCATCTCCGAAGTGAAACAAAGCAAGAGGATTTACTGTTTACAG 1040
QY 892 TATGATCATCAGAAATAGTAAAGGCGCTTCGAGACAAACACAAATCAAGATCTGGGATAA 951
DB 1041 TATGATCATCAGAAATAGTAAAGGCGCTTCGAGACAAACACAAATCAAGATCTGGGATAA 1100
QY 952 AACACATTTGAAATGCAAGCGAAATTTCTCAGAGGCCATACAGGTTTCACTCTCTCCAG 1011
DB 1101 AACACATTTGAAATGCAAGCGAAATTTCTCAGAGGCCATACAGGTTTCACTCTCTCCAG 1160
QY 1012 TATGATCAGAGATGATCATACAGGATCATCGGATTCACGGTTCAGAGTGGGATGTA 1071
DB 1161 TATGATCAGAGATGATCATACAGGATCATCGGATTCACGGTTCAGAGTGGGATGTA 1220
QY 1072 AATACAGGTGAAATGCTTAAACAGGTTGATTCACCATTTGTAAGCAGTTCCTGACCTTGGT 1131
DB 1221 AATACAGGTGAAATGCTTAAACAGGTTGATTCACCATTTGTAAGCAGTTCCTGACCTTGGT 1280
QY 1132 TTTCAATATGGCATGATGGTGACCTGCTCTCAAAGATCGTTCCATTCGTATGGGATATG 1191
DB 1281 TTTCAATATGGCATGATGGTGACCTGCTCTCAAAGATCGTTCCATTCGTATGGGATATG 1340
QY 1192 GCCTCCCAACTGACATTTACCTCCGAGGGTCTGCTCGGACACCGAGCTGCTGTCAAT 1251
DB 1341 GCCTCCCAACTGACATTTACCTCCGAGGGTCTGCTCGGACACCGAGCTGCTGTCAAT 1400
QY 1252 GTTGTAGACTTTGATGACAAAGTACATTTGTTCTGCAATCTGGGGATAGAACTATAAGGTA 1311
DB 1401 GTTGTAGACTTTGATGACAAAGTACATTTGTTCTGCAATCTGGGGATAGAACTATAAGGTA 1460
QY 1312 TGGAAACAAAGTACTCTGTAATTTGTAAGGACCTTAAATGGACACAAACGAGCAATGGC 1371
DB 1461 TGGAAACAAAGTACTCTGTAATTTGTAAGGACCTTAAATGGACACAAACGAGCAATGGC 1520
QY 1372 TGTTTGCACTACAGGACAGGCTGGTAGTGGTCTCATCTGACAACTATACAGTATA 1431
DB 1521 TGTTTGCACTACAGGACAGGCTGGTAGTGGTCTCATCTGACAACTATACAGTATA 1580
QY 1432 TGGGACATAGAAATGTTGTTGTCATGTTTACGAGTGTAGAGGCGCATGAGGAATTTGGTGGT 1491
DB 1581 TGGGACATAGAAATGTTGTTGTCATGTTTACGAGTGTAGAGGCGCATGAGGAATTTGGTGGT 1640
QY 1492 TGTTATTCGATTTGATACAAAGAGGATAGTCACTGGGGCTTATGATGGAAAAATTAAGTG 1551
DB 1641 TGTTATTCGATTTGATACAAAGAGGATAGTCACTGGGGCTTATGATGGAAAAATTAAGTG 1700
QY 1552 TGGGACATTTGTTGGTCTTTGGACCCCGTCTGCTCGAGGACACTCTGCTACGGAC 1611
DB 1701 TGGGACATTTGTTGGTCTTTGGACCCCGTCTGCTCGAGGACACTCTGCTACGGAC 1760
QY 1612 CTTGTGGAGCAATTCGGGAAGAGTTTTTCGACTACAGTTTTGATGAATTCAGATTTGTCAGT 1671
DB 1761 CTTGTGGAGCAATTCGGGAAGAGTTTTTCGACTACAGTTTTGATGAATTCAGATTTGTCAGT 1820
QY 1672 AGTTACATGATGACAACTCTCATCTGGGACTTCTAAATGATCCAGCTGCCAGCT 1731
DB 1821 AGTTACATGATGACAACTCTCATCTGGGACTTCTAAATGATCCAGCTGCCAGCT 1880
QY 1732 GAAACCCCGGTTCCCTCTCTCGAAACATACACCTACATCTCCAGATATAAATAACCATACAC 1791
DB 1881 GAAACCCCGGTTCCCTCTCTCGAAACATACACCTACATCTCCAGATATAAATAACCATACAC 1940
QY 1792 TGACCTCATACTTGGCCGAGGACCCATTTAAAG-TTGGCGTATTTTAAACGTATCTGCCAATAC 1850

```

Db      1941 TGACCTCATCTTGGCCAGGACCCATTAAAGTTTGGGGTATTTACGTATCTGCCATAC 2000
QY      1851 CAGGATGAGCAACACAGTAACAACTAACTACTGCCCCAGTTTCCCTGGAGTAGCCGAGG 1910
Db      2001 CAGGATGAGCAACACAGTAACAACTAACTACTGCCCCAGTTTCCCTGGAGTAGCCGAGG 2060
QY      1911 AGCAGGGCTTTGAGACTCTCTGTTGGGACACAGTTGGTCTGCAGTCGGCCAGGACGGTCT 1970
Db      2061 AGCAGGGCTTTGAGACTCTCTGTTGGGACACAGTTGGTCTGCAGTCGGCCAGGACGGTCT 2120
QY      1971 ACTAGCAACAATGACTGCTTCCAGTCGTGCTATACAGAAAGATGCTTCTATCAATTTGCAA 2030
Db      2121 ACTAGCAACAATGACTGCTTCCAGTCGTGCTATACAGAAAGATGCTTCTATCTTTGTGAA 2180
QY      2031 TGATTGGAATTTTAAACCTCCCTCCCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2090
Db      2181 TGATTGGAATTTTAAACCTCCCTCCCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2240
QY      2091 CCATTGGTCCAGACAAAGGTGACTTATAATATATTTAGTGTTTTGGCCGAA 2143
Db      2241 CCATTGGTCCAGACAAAGGTGACTTATAATATATTTAGTGTTTTGGCCGAA 2293

RESULT 11
AAK52699
ID      AAK52699 standard; cDNA; 3003 BP.
XX
AC      AAK52699;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human polynucleotide SEQ ID NO 2228.
XX
KW      Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW      vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW      tissue growth factor; immunomodulatory; cancer; leukaemia;
KW      nervous system disorder; arthritis; inflammation; ss.
XX
OS      Homo sapiens.
XX
FN      WO200157190-A2.
XX
PD      09-AUG-2001.
XX
PE      05-FEB-2001; 2001WO-US04098.
XX
PI      03-FEB-2000; 2000US-0496914.
PR      27-APR-2000; 2000US-0560875.
PR      20-JUN-2000; 2000US-0598075.
PR      19-JUL-2000; 2000US-0620325.
PR      01-SEP-2000; 2000US-0654936.
PR      15-SEP-2000; 2000US-0663561.
PR      20-OCT-2000; 2000US-0693325.
PR      30-NOV-2000; 2000US-0728422.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI      Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI      Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR      WPI; 2001-476283/51.
XX      P-PSDB; AAM79586.
XX
PT      Nucleic acids encoding polypeptides with cytokine-like activities,
PT      useful in diagnosis and gene therapy -
XX
PS      Claim 1; Page 4584-4595; 6221pp; English.
XX
CC      The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC      encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
CC      cytokine, cell proliferation or cell differentiation or which may induce

```

```

CC      production of other cytokines in other cell populations. The
CC      polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC      peptide therapy. The polypeptides have various cytokine-like activities,
CC      e.g. stem cell growth factor activity, haematopoiesis regulating
CC      activity, tissue growth factor activity, immunomodulatory activity and
CC      activin/inhibin activity and may be useful in the diagnosis and/or
CC      treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC      inflammation.
CC      Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC      (AAM80020) are omitted as the relevant pages from the sequence listing
CC      were missing at the time of publication.
XX
SQ      Sequence 3003 BP; 807 A; 703 C; 721 G; 772 T; 0 other;

Query Match
Best Local Similarity 93.4%; Score 2008.2; DB 22; Length 3003;
Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY      112 TTTATGAATTCCTCAGAGAGAGAGACTGTGTAATAATGGCGAACCCCTTAGGAGATAATA 171
Db      298 TTCAGAAATTCCTCAGAGAGAGAGAGACTGTGTAATAATGGCGAACCCCTTAGGAGATAATA 357
QY      172 CCAGAGAGAAATTCACCTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAACCAA 231
Db      358 CCAGAGAGAAATTCACCTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAACCAA 417
QY      232 GAAACAGTATGTTTAGCAAGCACTGCTATGAAGACTGAGAAATGTTGGCCAAAACAAA 291
Db      418 GAAACAGTATGTTTAGCAAGCACTGCTATGAAGACTGAGAAATGTTGGCCAAAACAAA 477
QY      292 CTTGCCAATGCACTTCCAGTATGATGTGCCCAAGCAACGGAACCTCTCAGCAGACTAT 351
Db      478 CTTGCCAATGCACTTCCAGTATGATGTGCCCAAGCAACGGAACCTCTCAGCAGACTAT 537
QY      352 GAAAGGAAAGGAACTGTGTCAATACTTTGAGCAGTGGTCAGAGTCAGATCAAGTG 411
Db      538 GAAAGGAAAGGAACTGTGTCAATACTTTGAGCAGTGGTCAGAGTCAGATCAAGTG 597
QY      412 GAATTTGTGGACATCTTATATCCAAAATGTCATTACCAACATGGGCACATAACTCG 471
Db      598 GAATTTGTGGACATCTTATATCCAAAATGTCATTACCAACATGGGCACATAACTCG 657
QY      472 TATCTTAAACCTATGTTGCAGAGAGATTTTCAATACCTGCTCTGCCAGCTCGGGGATGGAT 531
Db      658 TATCTTAAACCTATGTTGCAGAGAGATTTTCAATACCTGCTCTGCCAGCTCGGGGATGGAT 717
QY      532 CATATCGCTGAGAACATCTGTCTATACCTGGATGCCAAATCACTATGCTCTGTAACCT 591
Db      718 CATATTCCTGAGAACATCTGTCTATACCTGGATGCCAAATCACTATGCTCTGTAACCT 777
QY      592 GTGTGCAAGGAATGGTATCCGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCGAG 651
Db      778 GTGTGCAAGGAATGGTATCCGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCGAG 837
QY      652 AGAATGCTCAGACAGATTTCTGTGGAGAGGCTGSCAGAACCAAGAGATGGGGACAG 711
Db      838 AGAATGCTCAGACAGATTTCTGTGGAGAGGCTGSCAGAACCAAGAGATGGGGACAG 897
QY      712 TATTTATTTCAAAAAACAAACCTCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCA 771
Db      898 TATTTATTTCAAAAAACAAACCTCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCA 957
QY      772 CTTTATCTCTAAAATTATACAGACATTTAGACAAATAGAAATCTAATTTGGAGATGTGGAAGA 831
Db      958 CTTTATCTCTAAAATTATACAGACATTTAGACAAATAGAAATCTAATTTGGAGATGTGGAAGA 1017
QY      832 CATAGTTTTACAGAGAAATTCACCTGCGGAAGTGAACCAAGCAAGAGCTTTACTGTTTACAG 891
Db      1018 CATAGTTTTACAGAGAAATTCACCTGCGGAAGTGAACCAAGCAAGAGCTTTACTGTTTACAG 1077
QY      892 TATGATGATCAAAAAATAGTAAGCGGCTTTTCGAGACAAACAAATCAAGATCTGGGATAAA 951
Db      1078 TATGATGATCAAAAAATAGTAAGCGGCTTTTCGAGACAAACAAATCAAGATCTGGGATAAA 1137

```


Best Local Similarity 99.6%; Pred. No. 0;			
Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;			
Qy	112	TTTATGAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTTAGGAAGATAATA	171
Db	298	TTCCAGAAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTTAGGAAGATAATA	357
Qy	172	CCAGAGAAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTTAGGAAGATAATA	231
Db	358	CCAGAGAAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTTAGGAAGATAATA	417
Qy	232	GAACAGATGTTTACAGAGAGAACTGTAAATATGGCGAACCCCTTAGGAAGATAATA	291
Db	418	GAACAGATGTTTACAGAGAGAACTGTAAATATGGCGAACCCCTTAGGAAGATAATA	477
Qy	292	CTTGCCAAATGCTTCCAGATGATGTTGCGCAAGCAACGAACTCTCAGCAAGCTAT	351
Db	478	CTTGCCAAATGCTTCCAGATGATGTTGCGCAAGCAACGAACTCTCAGCAAGCTAT	537
Qy	352	GAAGAGAAAGAACTGTGTCAAAATACCTTTGAGCAAGTGTGAGAGTCAAGTCAAGTG	411
Db	538	GAAGAGAAAGAACTGTGTCAAAATACCTTTGAGCAAGTGTGAGAGTCAAGTCAAGTG	597
Qy	412	GAATTTGTGGAACATCTTATATCCCAAAATGTCATTTACCAACATGCGCAATAACTCG	471
Db	598	GAATTTGTGGAACATCTTATATCCCAAAATGTCATTTACCAACATGCGCAATAACTCG	657
Qy	472	TATCTTAAACCTATGTTGAGAGAGATTTTCAAACTGCTCTGCCAGCTCGGGGATTGGAT	531
Db	658	TATCTTAAACCTATGTTGAGAGAGATTTTCAAACTGCTCTGCCAGCTCGGGGATTGGAT	717
Qy	532	CATATCGCTCAGAACATCTGTATACCTGGATGGCAATCACTATGCTGCTGAACTT	591
Db	718	CATATGCTCAGAACATCTGTATACCTGGATGGCAATCACTATGCTGCTGAACTT	777
Qy	592	GTGTGCAAGAAATGTTACCGAGTACCTCTGTATGGCATGCTGTGGAAGAACTTATCGAG	651
Db	778	GTGTGCAAGAAATGTTACCGAGTACCTCTGTATGGCATGCTGTGGAAGAACTTATCGAG	837
Qy	652	AGATGCTCAGGACAGATTTCTGTGGAGAGCCCTGGCAGAACGAGAGATGGGACAG	711
Db	838	AGATGCTCAGGACAGATTTCTGTGGAGAGCCCTGGCAGAACGAGAGATGGGACAG	897
Qy	712	TATTTATCAAAACCAAACTCTGTGCGGAAATGCTCTCCCAACTCTTTTATAGAGCA	771
Db	898	TATTTATCAAAACCAAACTCTGTGCGGAAATGCTCTCCCAACTCTTTTATAGAGCA	957
Qy	772	TTTATCTTAAATATACAGACATGTAGCAATAGATCTAATTTGAGATGTTGGAAGA	831
Db	958	TTTATCTTAAATATACAGACATGTAGCAATAGATCTAATTTGAGATGTTGGAAGA	1017
Qy	832	CATAGTTTACAGAGAAATTCATCTCCGAAAGTGAACCAAGCAAGAGAGTTTACTGTTTACAG	891
Db	1018	CATAGTTTACAGAGAAATTCATCTCCGAAAGTGAACCAAGCAAGAGAGTTTACTGTTTACAG	1077
Qy	892	TATGATCATCAGAAATAGTAAGCGCCCTTCGAGACACACATCAAGATCTGGATATAA	951
Db	1078	TATGATCATCAGAAATAGTAAGCGCCCTTCGAGACACACATCAAGATCTGGATATAA	1137
Qy	952	AACACATTGGAATGCAAGCGAATTTTACAGAGCCATACAGGTTTCTGCTCTCCAG	1011
Db	1138	AACACATTGGAATGCAAGCGAATTTTACAGAGCCATACAGGTTTCTGCTCTCCAG	1197
Qy	1012	TATGATCAGAGATGATCATTAACAGATCATCGATTTCCAGGTCAGAGTGGGATGTA	1071
Db	1198	TATGATCAGAGATGATCATTAACAGATCATCGATTTCCAGGTCAGAGTGGGATGTA	1257
Qy	1172	AATACAGTGAAATGCTTAAACAGCTTGAATTCACATTTGTAAGCAGTTCTGCACCTTGGT	1131
Db	1258	AATACAGTGAAATGCTTAAACAGCTTGAATTCACATTTGTAAGCAGTTCTGCACCTTGGT	1317
Qy	1132	TTCAATATGGCATGATGGTACCTGCTCCAAAGATCGTTCCATGCTGTATGGGATATG	1191

Db	1318	TTCAATATGGCATGATGGTGAACCTCTCCAAAGATGTTCCATCTGCTGTATGGGATATG	1377
Qy	1192	GCCTCCCAACTGACATTTACCTCCGAGAGGTGCTGGTCGACACCGAGCTGCTGCTCAAT	1251
Db	1378	GCCTCCCAACTGACATTTACCTCCGAGAGGTGCTGGTCGACACCGAGCTGCTGCTCAAT	1437
Qy	1252	GTTGTAGACTTTGATGACAAAGTACATTTGTTCTGCACTGGGGATAGAACTATAAAGGTA	1311
Db	1438	GTTGTAGACTTTGATGACAAAGTACATTTGTTCTGCACTGGGGATAGAACTATAAAGGTA	1497
Qy	1312	TGGAACCAAAAGTACTTGTGAAATTTGTAAGGACCTTAAATGGAACAAACGAGGCAATGCC	1371
Db	1498	TGGAACCAAAAGTACTTGTGAAATTTGTAAGGACCTTAAATGGAACAAACGAGGCAATGCC	1557
Qy	1372	TGTTTGCAGTACAGGACAGGCTGGTAGTGAGTGGCTCATCTGCAACACACTATCAGATTA	1431
Db	1558	TGTTTGCAGTACAGGACAGGCTGGTAGTGAGTGGCTCATCTGCAACACACTATCAGATTA	1617
Qy	1432	TGGACATAGAAATGTGGTGCATGTTTACGAGTGTAGAGGCCATGAGGAAATTTGGTGGGT	1491
Db	1618	TGGACATAGAAATGTGGTGCATGTTTACGAGTGTAGAGGCCATGAGGAAATTTGGTGGGT	1677
Qy	1492	TGTTTGCAGTACAGGACAGGATAGTACGAGTGGGCTTATGATGGAATAAATAAGTG	1551
Db	1678	TGTTTGCAGTACAGGACAGGATAGTACGAGTGGGCTTATGATGGAATAAATAAGTG	1737
Qy	1552	TGGGATCTTGTGGCTGCTTTTGGACCCCGTGTCTTGCAGGACACTCTGTCTACGGACC	1611
Db	1738	TGGGATCTTGTGGCTGCTTTTGGACCCCGTGTCTTGCAGGACACTCTGTCTACGGACC	1797
Qy	1612	CTTGTGAGCATTTCCGGAAGAGTTTTCGACATACAGTTTGTAGTAATTCAGATTTGTCAGT	1671
Db	1798	CTTGTGAGCATTTCCGGAAGAGTTTTCGACATACAGTTTGTAGTAATTCAGATTTGTCAGT	1857
Qy	1672	AGTTTACATGATGACAAATCTCTATCTGGGATCTCTTAAATGATCCAGTGGCCCAAGCT	1731
Db	1858	AGTTTACATGATGACAAATCTCTATCTGGGATCTCTTAAATGATCCAGTGGCCCAAGCT	1917
Qy	1732	GAACCCCGGTTTCCCTTCTCGAACATACACTACATCTCCAGATAAATACCATACAC	1791
Db	1918	GAACCCCGGTTTCCCTTCTCGAACATACACTACATCTCCAGATAAATACCATACAC	1977
Qy	1792	TGACCTCATACTTGGCCAGGACCCATTTAAAGTTTGGGATTTTAAACGATCTGCCAATAC	1850
Db	1978	TGACCTCATACTTGGCCAGGACCCATTTAAAGTTTGGGATTTTAAACGATCTGCCAATAC	2037
Qy	1851	CAGGATGAGCAACAAAGTAAACAACTCAAACTACTGCCAGTTTCCCTGAGCTAGCCGAGG	1910
Db	2038	CAGGATGAGCAACAAAGTAAACAACTCAAACTACTGCCAGTTTCCCTGAGCTAGCCGAGG	2097
Qy	1911	AGCAGGCTTTGAGACTCTTGTGGACACAGTTGGTCTGAGCTCGGCCAGGACGGTCT	1970
Db	2098	AGCAGGCTTTGAGACTCTTGTGGACACAGTTGGTCTGAGCTCGGCCAGGACGGTCT	2157
Qy	1971	ACTCAGACAACTGACGCTTCACTGCTGATCAAGAGATGCTCTTCTATCAATTTGTA	2030
Db	2158	ACTCAGACAACTGACGCTTCACTGCTGATCAAGAGATGCTCTTCTATCAATTTGTA	2217
Qy	2031	TGATTTGGAATTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2090
Db	2218	TGATTTGGAATTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2277
Qy	2091	CCATTTGGTTCAGACAAAGGTGACTTATAAATATATTTTAGTGTGTTTGGCAGAA	2143
Db	2278	CCATTTGGTTCAGACAAAGGTGACTTATAAATATATTTTAGTGTGTTTGGCAGAA	2330

RESULT 13
AAK52701
ID AAK52701 standard; cdna; 3003 BP.
XX
AC AAK52701;
XX

JT 06-MQV-2001 (first entry)
KW Human polynucleotide SEQ ID NO 2230.
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
KW Homo sapiens.
KW WO200157190-A2.
KW 09-AUG-2001.
KW 05-FEB-2001; 2001WO-US04098.
KW 03-FEB-2000; 2000US-0496914.
KW 27-APR-2000; 2000US-0560875.
KW 20-JUN-2000; 2000US-0598075.
KW 19-JUL-2000; 2000US-0620325.
KW 01-SEP-2000; 2000US-0654936.
KW 15-SEP-2000; 2000US-0663561.
KW 20-OCT-2000; 2000US-0693325.
KW 30-NOV-2000; 2000US-0728422.
KW (HYSB-) HYSEQ INC.
KW Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
KW Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
KW Xue AJ, Yang Y, Wejhrman T, Goodrich R;
KW MPI; 2001-476283/51.
KW P-PSDB; AAM79568.
KW Nucleic acids encoding polypeptides with cytokine-like activities,
KW useful in diagnosis and gene therapy -
KW Claim 1; Page 4586-4587; 6221pp; English.
KW The invention relates to polynucleotides (AAK51456-AAK53435) and the
KW encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
KW cytokines, cell proliferation or cell differentiation or which may induce
KW production of other cytokines in other cell populations. The
KW polynucleotides and polypeptides are useful in gene therapy, vaccines or
KW peptide therapy. The polypeptides have various cytokine-like activities,
KW e.g. stem cell growth factor activity, haematopoiesis regulating
KW activity, tissue growth factor activity, immunomodulatory activity and
KW activin/inhibin activity and may be useful in the diagnosis and/or
KW treatment of cancer, leukaemia, nervous system disorders, arthritis and
KW inflammation.
KW Note: Records for SEQ ID NO 2110 (AAK52591), 2111 (AAK52582) and 3666
KW (AAM80020) are omitted as the relevant pages from the sequence listing
KW were missing at the time of publication.
KW Sequence 3003 BP; 807 A; 703 C; 721 G; 772 T; 0 other;
Query Match 93.4%; Score 2008.2; DB 22; Length 3003;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
112 TTATGAATTCCTCAGAGAGAGAGACTGTAATATGGCGAACCCCTAGGAAGATAATA 171
172 CCAGAGAGAGATTCATTAGACAGACATACACAGCTGTGCCAGACTCTGTTAAACCAA 231
358 CCAGAGAGAGATTCATTAGACAGACATACACAGCTGTGCCAGACTCTGTTAAACCAA 417
232 GAAACAGTAGTTTACAGAGCACTGCTATGAAGACTGAGAAATTTGTGGCCCAACAAA 291
418 GAAACAGTAGTTTACAGAGCACTGCTATGAAGACTGAGAAATTTGTGGCCCAACAAA 477

QY 292 CTTGCCAATGSCACTTCCRGATATGATGTGCCCAAGCAACGGAAACTCTCAGCAAGCTAT 351
DB 478 CTTGCCAATGSCACTTCCAGTATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTAT 537
QY 352 GAAAGGAAAGGAACTGTGTGCAATACTTTGAGCAGTGGTTCAGAGTCAGATCAAGTG 411
DB 538 GAAAGGAAAGGAACTGTGTGCAATACTTTGAGCAGTGGTTCAGAGTCAGATCAAGTG 597
QY 412 GAATTTGTGGAACATCTTATATATCCAAATGTGTCATTACCAACATGGGCACATAAATCG 471
DB 598 GAATTTGTGGAACATCTTATATCCAAATGTGTCATTACCAACATGGGCACATAAATCG 657
QY 472 TATCTTAACCTATGTTGCAGAGAGATTTTCAATCTCTCCAGCTCGGGATGGAT 531
DB 658 TATCTTAACCTATGTTGCAGAGAGATTTTCAATCTCTCCAGCTCGGGATGGAT 717
QY 532 CATATCGCTGAGAACATTTCTGTATACCTGGATGCCAAATCACTATGTCTGCTGAACAT 591
DB 718 CATATGCTGAGAACATTTCTGTATACCTGGATGCCAAATCACTATGTCTGCTGAACAT 777
QY 592 GTGTGCAAGGAAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAG 651
DB 778 GTGTGCAAGGAAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAG 837
QY 652 AGAATGCTCAGGACAGATTTCTGTGAGAGGCTGGCAGAACGAAAGAGAGATGGGACAG 711
DB 838 AGAATGCTCAGGACAGATTTCTGTGAGAGGCTGGCAGAACGAAAGAGAGATGGGACAG 897
QY 712 TATTTATTCAAACAAACCTCTCTGACGGGAATGCTCTCCCACTCTTTTATAGAGCA 771
DB 898 TATTTATTCAAACAAACCTCTCTGACGGGAATGCTCTCCCACTCTTTTATAGAGCA 957
QY 772 CTTTATCTCTAAAATTATACAGACATTTGACAAATAGAATCTAATTTGGAGATGTGGAAGA 831
DB 958 CTTTATCTCTAAAATTATACAGACATTTGACAAATAGAATCTAATTTGGAGATGTGGAAGA 1017
QY 832 CATAGTTTACAGAGAAATTCATGCCGAGTGAACAGCAAGAGAGTTTACTGTTTACAG 891
DB 1018 CATAGTTTACAGAGAAATTCATGCCGAGTGAACAGCAAGAGAGTTTACTGTTTACAG 1077
QY 892 TATCATGATCAGAAATAGTAGCGGCTTTTCGAGACAAACACAAATCAAGATCTTGGGATAAA 951
DB 1078 TATCATGATCAGAAATAGTAGTAGCGGCTTTTCGAGACAAACACAAATCAAGATCTTGGGATAAA 1137
QY 952 AACCATTTGGAATCCAAAGCGAAATTTTCACAGGCCATACAGGTTTCACTGTTCTCCAG 1011
DB 1138 AACCATTTGGAATCCAAAGCGAAATTTTCACAGGCCATACAGGTTTCACTGTTCTCCAG 1197
QY 1012 TATGATGAGAGAGTGATCATAACAGGATCATCGGATTCACCGTTCAGAGTGTGGATGTA 1071
DB 1198 TATGATGAGAGAGTGATCATAACAGGATCATCGGATTCACCGTTCAGAGTGTGGATGTA 1257
QY 1072 AATCAGGTGMAATGCTAAACAGCTTGATTCACCATTTGAGAGAGTTCTGCACTTGGCT 1131
DB 1258 AATCAGGTGMAATGCTAAACAGCTTGATTCACCATTTGAGAGAGTTCTGCACTTGGCT 1317
QY 1132 TTCATATGTCATGATGTCGCTGCTCCAAAGATGCTTCCATTCCTGATGGGATATG 1191
DB 1318 TTCATATGTCATGATGTCGCTGCTCCAAAGATGCTTCCATTCCTGATGGGATATG 1377
QY 1192 GCCTCCCCCACTGACATTTACCTCCGAGGGTCTGGTGGACACCCAGCTGCTGTCAAT 1251
DB 1378 GCCTCCCCCACTGACATTTACCTCCGAGGGTCTGGTGGACACCCAGCTGCTGTCAAT 1437
QY 1252 GTTGTAGACTTTGATGACAGTACATGTTTCTGCACTTGGGATAGACTATAGGTA 1311
DB 1438 GTTGTAGACTTTGATGACAGTACATGTTTCTGCACTTGGGATAGACTATAGGTA 1497
QY 1312 TGGAAACACAGTACTTGTGAATTTGTAAAGGACCTTAAATGGACACAAACGAGGCAATGCG 1371
DB 1498 TGGAAACACAGTACTTGTGAATTTGTAAAGGACCTTAAATGGACACAAACGAGGCAATGCG 1557
QY 1372 TGTGTCAGATACAGGACAGCTGGTAGTGCTGCTGACAACTATCAGATTA 1431

601	TTCTCTGTGGAGAGCCCTTGGCAGAACGAGAGGATGGGACAGTATTTTATTCAAAACAAA	660
730	CCTCCTCAGCGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATTCCTAAAATTATATA	789
661	CCTCCTGACGGGAATGCTCCTCCCACTCTTTTATAGAGCACTTTATTCCTAANAATTATA	720
790	CAAGACATTTGAGACAAATAGAAATCTAATTTGGAGATGTGGAAGACATPAGTTTACAGAGAATT	849
721	CAAGACATTTGAGACAAATAGAAATCTAATTTGGAGATGTGGAAGACATPAGTTTACAGAGAATT	780
850	CACCTCCCAAGTGTAAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGATCAGAAAATA	909
781	CACTGCCGAAGTGAAACAGCAAAAGAGTTTACTGTTTACAGTATGATGATCAGAAAATA	840
910	GTAAAGCGCCCTTCGAGACAAACAACATCAAGATCTGGATATAAAACACATTCGAAATGCAAG	969
841	GTAAAGCGCCCTTCGAGACAAACAACATCAAGATCTGGATATAAAACACATTCGAAATGCAAG	900
970	CGAATTTCTCAGAGCCCATACAGGTTTCAGTCTCTGCTCCAGTATGATGAGAGAGTCAATC	1029
901	CGAATTTCTCAGAGCCCATACAGGTTTCAGTCTCTGCTCCAGTATGATGAGAGAGTCAATC	960
1030	ATAACAGGATCATCGGAATCCACGGTCAGAGTGTGGGATGTAAATACAGGTGAAATGCTTA	1089
961	ATAACAGGATCATCGGAATCCACGGTCAGAGTGTGGGATGTAAATACAGGTGAAATGCTTA	1020
1090	AACAGGTTTGATTCACCATTTGGAAGCAGTTCTGCATCTCGGTTTCAATTAATGGCATGATG	1149
1021	AACAGGTTTGATTCACCATTTGGAAGCAGTTCTGCATCTCGGTTTCAATTAATGGCATGATG	1080
1150	GTGACCTCTCTCCAAAGATTCGTTCCATTTGCTGTATGGGATATGGCCTCCCCAACTGACATT	1209
1081	GTGACCTCTCTCCAAAGATTCGTTCCATTTGCTGTATGGGATATGGCCTCCCCAACTGACATT	1140
1210	ACCCCTCCGAGGGTGCTGCTGCGACACCGAGCTGCTGTCAAAGTTGTAGACTTTTGATCAC	1269
1141	ACCCCTCCGAGGGTGCTGCTGCGACACCGAGCTGCTGTCAAAGTTGTAGACTTTTGATCAC	1200
1270	AAGTACATTTGTTCTGCATCTGGGGATAGAACTATAAAGGTATGGAAACAAGTACTTGT	1329
1201	AAGTACATTTGTTCTGCATCTGGGGATAGAACTATAAAGGTATGGAAACAAGTACTTGT	1260
1330	GAAATTTGTAAGACCTTAAATTGGACACAAACGAGGCATTTGCCCTGTTTGGCATACAGGGAC	1389
1261	GAAATTTGTAAGACCTTAAATTGGACACAAACGAGGCATTTGCCCTGTTTGGCATACAGGGAC	1320
1390	AGGCTGGTAGTGGAGTGCTCATCTGACAAACACTATCAGATTAATGGACATAGAATTGGT	1449
1321	AGGCTGGTAGTGGAGTGCTCATCTGACAAACACTATCAGATTAATGGACATAGAATTGGT	1380
1450	GCATGTTTACAGATGTTTAGAAGGCCATAGGAATTCGTCGGTTGTATTCGATTTGATAAC	1509
1381	GCATGTTTACAGATGTTTAGAAGGCCATAGGAATTCGTCGGTTGTATTCGATTTGATAAC	1440
1510	AAGAGGATAGTCAGTGGGGCCTATGATGGGAATTAAGATCTGGGATCTTGTGGCTGCT	1569
1441	AAGAGGATAGTCAGTGGGGCCTATGATGGGAATTAAGATCTGGGATCTTGTGGCTGCT	1500
1570	TTTGACCCCGCTGCTCTCGACGGGACACTCTGTCTACGGACCCCTTGTGGAGCATTTCCGGA	1629
1501	TTTGACCCCGCTGCTCTCGACGGGACACTCTGTCTACGGACCCCTTGTGGAGCATTTCCGGA	1560
1630	AGAGTTTTTCGACTACAGTTTGATGAATTCGAAATTCGAGTTGTCAAGTAGTGACACACA	1689
1561	AGAGTTTTTCGACTACAGTTTGATGAATTCGAAATTCGAGTTGTCAAGTAGTGACACACA	1620
1690	ATTCCTCATCTGGGACTCTCCTAAATGATCCAGCTGCCCAAGCTGAACCCCGCTTCCCT	1749
1621	ATTCCTCATCTGGGACTCTCCTAAATGATCCAGCTGCCCAAGCTGAACCCCGCTTCCCT	1680
1750	TTCTCGAACAATACACCTACATCTCCAGA	1776

Db 1681 TTCGACATACACCTACATCTCCAGA 1707

RESULT 15
AAKS1715
ID AAKS1715 standard; cDNA; 2285 BP.
XX
XX AAKS1715;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 260.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US040398.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.
XX 01-SEP-2000; 2000US-0654936.
XX 15-SEP-2000; 2000US-0663561.
XX 20-OCT-2000; 2000US-0693325.
XX 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang J, Ren P, Chen R, Wang ZW;
XX Xie AJ, Wang Y, Wejhrman T, Goodrich R.
XX
XX WPI; 2001-476283/51.
XX P-PSDB; AAM78582.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX
XX Claim 1; Page 1177-1180; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAKS1456-AAKS3435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX
XX Note: Records for SEQ ID NO 2110 (AAKS2581), 2111 (AAKS2582) and 3666
XX (AAM90020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
XX Sequence 2285 BP; 646 A; 492 C; 573 G; 574 T; 0 other;

2y 1730 TC1CGAACATACACCTACATCTCCAGH 1776

QY	172	CCAGAGAAGAAATTCATTAGACAGACATACAAACAGCTGTGCCAGACTCTGCTTAAACCAA	231
Db	337	CCAGAGAAGAAATTCGCTTAGACAGACATACAAACAGCTGTGCCAGACTCTGCTTAAACCAA	396
QY	232	GAACAGTAGTGTAGACAGACTCTGTATGAGACTGAGATTCTGTGGCCAAACAA	291
Db	397	GAACAGTAGTGTAGACAGACTCTGTATGAGACTGAGATTCTGTGGCCAAACAA	456
QY	292	CTTGCCAAATGGCACTTCCAGTAGTGTGGCCAAAGCAACGAACTCTCAGCAAGCTAT	351
Db	457	CTTGCCAAATGGCACTTCCAGTAGTGTGGCCAAAGCAACGAACTCTCAGCAAGCTAT	516
QY	352	GAACAGTAGTGTAGACAGACTCTGTATGAGACTGAGATTCTGTGGCCAAACAA	411
Db	517	GAACAGTAGTGTAGACAGACTCTGTATGAGACTGAGATTCTGTGGCCAAACAA	576
QY	412	GAATTTGTGGAACATCTTATATCCCAAAATGTGTCAATACCAACATGGGCAATAACTCG	471
Db	577	GAATTTGTGGAACATCTTATATCCCAAAATGTGTCAATACCAACATGGGCAATAACTCG	636
QY	472	TATCTTAAACCTATGTGACAGAGATTTCATAAGCTCTGCCAGCTCGGGGATTGGAT	531
Db	637	TATCTTAAACCTATGTGACAGAGATTTCATAAGCTCTGCCAGCTCGGGGATTGGAT	696
QY	532	CATATCGCTGAGAACATCTGTATACCTGGATGCCAAATCACTATGTGTCTGAACTT	591
Db	697	CATATCGCTGAGAACATCTGTATACCTGGATGCCAAATCACTATGTGTCTGAACTT	756
QY	592	GTGTGAAGAAATGATGACAGAGACTCTGTATGAGACTGAGATTCTGTGGCCAAACAA	651
Db	757	GTGTGAAGAAATGATGACAGAGACTCTGTATGAGACTGAGATTCTGTGGCCAAACAA	816
QY	652	AGAATGCTCAGGACAGATTCTCTGTGGAGAGCCCTGGCAGAACGAAAGGATGGGACAG	711
Db	817	AGAATGCTCAGGACAGATTCTCTGTGGAGAGCCCTGGCAGAACGAAAGGATGGGACAG	876
QY	712	TATTTATTCAAACAAACCTCTGTGAGAGAGCTCTGTGAGAGAGCTCTCTTATAGAGCA	771
Db	877	TATTTATTCAAACAAACCTCTGTGAGAGAGCTCTGTGAGAGAGCTCTCTTATAGAGCA	936
QY	772	CTTTATCTCTAAATATACAGACATTTAGACAAATAGAAATCTAAATGGAGATGGGAAGA	831
Db	937	CTTTATCTCTAAATATACAGACATTTAGACAAATAGAAATCTAAATGGAGATGGGAAGA	996
QY	832	CATAGTTTACAGAGAAATTCACCTGCCAGTGAACAGCAAGGAGTTTACTGTTTACAG	891
Db	997	CATAGTTTACAGAGAAATTCACCTGCCAGTGAACAGCAAGGAGTTTACTGTTTACAG	1056
QY	892	TATGATGATCAGAAATAGTAAGCGCCCTTCGAGACAAACAAATCAGATCTGGGATAAA	951
Db	1057	TATGATGATCAGAAATAGTAAGCGCCCTTCGAGACAAACAAATCAGATCTGGGATAAA	1116
QY	952	AACACATTTGAATGCAAGCGAATCTTCACAGCCCATACAGGTTTCACTCTCTCTCCAG	1011
Db	1117	AACACATTTGAATGCAAGCGAATCTTCACAGCCCATACAGGTTTCACTCTCTCTCCAG	1176
QY	1012	TATGATGAGAGATGATCATACAGGATCATCGGATTCACGGTTCAGAGTGTGGATGTA	1071
Db	1177	TATGATGAGAGATGATCATACAGGATCATCGGATTCACGGTTCAGAGTGTGGATGTA	1236
QY	1072	AATACAGGTGAAATGCTAAACAGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTGGT	1131
Db	1237	AATACAGGTGAAATGCTAAACAGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTGGT	1296
QY	1132	TTCAATTAATGGCATGATGTTGACCTCTCCAAAGATCGTTCCATTTGTTATGGGATATG	1191
Db	1297	TTCAATTAATGGCATGATGTTGACCTCTCCAAAGATCGTTCCATTTGTTATGGGATATG	1356
QY	1192	GCCTCCCAACTGACATTTACCTCCGAGGGTCTGTGCGACCCGAGCTCTGTCTCAAT	1251
Db	1357	GCCTCCCAACTGACATTTACCTCCGAGGGTCTGTGCGACCCGAGCTCTGTCTCAAT	1416

QY	1252	GTGTAGACTTTTATGACAAAGTACATTTGTTTCTGCACTGTGGGATAGAACTATATAAGTA	1311
Db	1417	GTGTAGACTTTTATGACAAAGTACATTTGTTTCTGCACTGTGGGATAGAACTATATAAGTA	1476
QY	1312	TGGAACACAAAGTACTTGTGAAATTTGTAAGGACTTTAAATGGACACAAACGAGGCAATTGCC	1371
Db	1477	TGGAACACAAAGTACTTGTGAAATTTGTAAGGACTTTAAATGGACACAAACGAGGCAATTGCC	1536
QY	1372	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGACACACACTATCAGATTA	1431
Db	1537	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGACACACACTATCAGATTA	1596
QY	1432	TGGGACATAGAAATGTGTGCACTTTTACAGTGTGTAGAGGCCCATGAGAAATTTGTCGT	1491
Db	1597	TGGGACATAGAAATGTGTGCACTTTTACAGTGTGTAGAGGCCCATGAGAAATTTGTCGT	1656
QY	1492	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGACACACACTATCAGATTA	1551
Db	1657	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGACACACACTATCAGATTA	1716
QY	1552	TGGGATCTTGTGCTGCTTTGGACCCCGTGTCTGACAGGACACTCTGTCTACGGACC	1611
Db	1717	TGGGATCTTGTGCTGCTTTGGACCCCGTGTCTGACAGGACACTCTGTCTACGGACC	1776
QY	1612	CTTGTGAGGCAATTCGGGAGAGTTTTCGACTACAGTTTGTAGTGAATTCAGATTTGTCAGT	1671
Db	1777	CTTGTGAGGCAATTCGGGAGAGTTTTCGACTACAGTTTGTAGTGAATTCAGATTTGTCAGT	1836
QY	1672	AGTTTACATGATGACACAACTCTCTGAGACTTCTTAAATGATCCAGCTGCCCAAGCT	1731
Db	1837	AGTTTACATGATGACACAACTCTCTGAGACTTCTTAAATGATCCAGCTGCCCAAGCT	1896
QY	1732	GAACCCCGGTTTCCCGTTCTCGAACATACCTACATCTCCAGATTAATACCATACAC	1791
Db	1897	GAACCCCGGTTTCCCGTTCTCGAACATACCTACATCTCCAGATTAATACCATACAC	1956
QY	1792	TGACCTCATACTTTGCCCGAGCCCATTTAAAGTTGCGGTATTTTAAAC	1836
Db	1957	TGACCTCATACTTTGCCCGAGGTATCGAAATCGATATGTACATAAC	2001

Search completed: October 22, 2003, 18:40:07
Job time : 567 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

XM nucleic - nucleic search, using sw model

run on: October 22, 2003, 20:49:00 ; Search time 576 Seconds
(without alignments)
10014.850 Million cell updates/sec

Title: US-09-601-168B-1

perfect score: 2151
sequence: 1 tgcgttgctgcgcctggc.....gtttgcagaaaaaaa 2151

scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 134090451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2151	100.0	2151	12	US-09-832-161-17 Sequence 17, Appl
2	2151	100.0	2151	13	US-10-042-417-1 Sequence 1, Appl
3	1710	79.5	1710	14	US-10-038-010-7 Sequence 7, Appl
4	883.4	41.1	4230	12	US-09-832-161-15 Sequence 15, Appl
5	657	30.5	657	14	US-10-023-530-1 Sequence 1, Appl
6	435.4	20.2	951	9	US-09-764-848-15 Sequence 15, Appl
7	435.4	20.2	951	12	US-10-222-020-15 Sequence 15, Appl
8	435.4	20.2	951	14	US-10-116-016-15 Sequence 15, Appl
9	272	12.6	479	9	US-09-864-761-2778 Sequence 15, Appl
10	208.4	9.7	370	9	US-09-864-761-19492 Sequence 15, Appl
11	180.2	8.4	319	10	US-09-960-352-2934 Sequence 2934, Ap
12	167.8	7.8	524	12	US-10-029-386-4953 Sequence 4953, Ap
13	162.6	7.6	250	12	US-10-029-386-18709 Sequence 18709, A
14	157.4	7.3	475	11	US-09-948-995-15756 Sequence 15756, A
15	124.2	5.8	1620	12	US-10-245-618-13 Sequence 13, Appl
16	124.2	5.8	1659	12	US-10-245-618-7 Sequence 7, Appl

17	124.2	5.8	1770	12	US-10-245-618-17 Sequence 17, Appl
18	124.2	5.8	1881	10	US-09-213-888-20 Sequence 20, Appl
19	124.2	5.8	1881	10	US-09-328-877A-20 Sequence 20, Appl
20	124.2	5.8	1884	12	US-10-245-618-5 Sequence 5, Appl
21	124.2	5.8	2001	10	US-09-213-888-26 Sequence 26, Appl
22	124.2	5.8	2001	10	US-09-328-877A-26 Sequence 26, Appl
23	124.2	5.8	2010	10	US-09-213-888-24 Sequence 24, Appl
24	124.2	5.8	2010	10	US-09-328-877A-24 Sequence 24, Appl
25	124.2	5.8	2063	12	US-10-245-618-34 Sequence 34, Appl
26	124.2	5.8	2124	12	US-10-245-618-1 Sequence 1, Appl
27	124.2	5.8	2255	12	US-10-245-618-35 Sequence 35, Appl
28	124.2	5.8	3550	10	US-09-213-888-1 Sequence 1, Appl
29	124.2	5.8	3550	10	US-09-328-877A-1 Sequence 1, Appl
30	124.2	5.8	3571	10	US-09-213-888-2 Sequence 2, Appl
31	124.2	5.8	3571	10	US-09-328-877A-2 Sequence 2, Appl
32	82.2	3.8	2237	10	US-09-994-485-7 Sequence 7, Appl
33	82.2	3.8	2237	10	US-09-832-292-11 Sequence 11, Appl
34	71.8	3.3	2208	12	US-10-032-585-6658 Sequence 6658, Ap
35	62.4	2.9	2091	14	US-10-128-714-2506 Sequence 2506, Ap
36	62.4	2.9	2091	14	US-10-128-714-2506 Sequence 2506, Ap
37	62.4	2.9	2167	14	US-10-128-714-1506 Sequence 1506, Ap
38	62.4	2.9	2167	14	US-10-128-714-1506 Sequence 1506, Ap
39	62.4	2.9	4167	14	US-10-128-714-506 Sequence 506, App
40	62.4	2.9	4167	14	US-10-128-714-506 Sequence 506, App
41	60	2.8	60	12	US-09-908-975-14380 Sequence 14380, A
42	57	2.6	65	12	US-09-908-975-29374 Sequence 29374, A
43	55.2	2.6	2037	10	US-09-801-368-313 Sequence 313, App
44	53.8	2.5	475	11	US-09-918-995-19598 Sequence 19598, A
45	53.8	2.5	10668	12	US-09-814-353-20029 Sequence 20029, A

ALIGNMENTS

RESULT 1

US-09-832-161-17
; Sequence 17, Application US/09832161
; Publication No. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercutio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinson
; APPLICANT: Davis, Matti
; APPLICANT: Hatzubai, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: NF-KB
; FILE REFERENCE: 860098.427
; CURRENT APPLICATION NUMBER: US/09/832,161
; CURRENT FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US/210,060
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO. 17
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-161-17

Query Match 100.0%; Score 2151; DB 12; Length 2151;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCTTGGCTGGCGCTGGACCAAGGGCGCCCGCGGAGAGCGAGCCAGTGGCC 60

Db 1 TCGCTTGGCTGGCGCTGGACCAAGGGCGCCCGCGGAGAGCGAGCCAGTGGCC 60

QY 61 TCGCGGATTATGCGCCCGCGGCGGGGCTCTCAGAGAGAGCGCACTCAAGTTATGAAT 120

Db 61 TCGCGGATTATGCGCCCGCGGCGGGGCTCTCAGAGAGAGCGCACTCAAGTTATGAAT 120


```

; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-10-042-417-1

Query Match      100.0%; Score 2151; DB 13; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 TCGTTGGCTGGCGCTGGCCACCAAGAGGGCGGCCCGCGGAGAGGGAGCCCAAGTGGCC 60
Db 1 TCGTTGGCTGGCGCTGGCCACCAAGAGGGCGGCCCGCGGAGAGGGAGCCCAAGTGGCC 60

2Y 61 TCGGCGATTATGGACCCCGCCGAGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAAT 120
Db 61 TCGGCGATTATGACCCCGCCGAGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAAT 120

2Y 121 TCTCAGAGAGAGAGACTGTAAATATGGCGAAACCCCTAGGAAGATAATACCAAGAGAAG 180
Db 121 TCTCAGAGAGAGAGACTGTAAATATGGCGAAACCCCTAGGAAGATAATACCAAGAGAAG 180

2Y 181 AATTCACTTAGACAGACATACAAAGCTGTGCGAGACTCTGCTTAAACCAAGAAACAGTA 240
Db 181 AATTCACTTAGACAGACATACAAAGCTGTGCGAGACTCTGCTTAAACCAAGAAACAGTA 240

2Y 241 TGTTTAGCAAGCACTGCTATGAAGACTGAGAAATGTGTGGCCAAACAAACTTGCCTAAT 300
Db 241 TGTTTAGCAAGCACTGCTATGAAGACTGAGAAATGTGTGGCCAAACAAACTTGCCTAAT 300

2Y 301 GGCACCTTCAGTATGATTGTGCGCAAGCAACGGAACCTCTCAGGAGCTATGAAAGAGAA 360
Db 301 GGCACCTTCAGTATGATTGTGCGCAAGCAACGGAACCTCTCAGGAGCTATGAAAGAGAA 360

2Y 361 AAGGAACCTGTGTCAATPACTTTGAGCAGTGGTCAGAGTCAAGTCAAGTGGAAATTTGTG 420
Db 361 AAGGAACCTGTGTCAATPACTTTGAGCAGTGGTCAGAGTCAAGTCAAGTGGAAATTTGTG 420

2Y 421 GAACATCTTATATCCCAATGTGTCAATACCAATGGGCAATAAATCTGTATCTTAAA 480
Db 421 GAACATCTTATATCCCAATGTGTCAATACCAATGGGCAATAAATCTGTATCTTAAA 480

2Y 481 CCTATGTCAGAGAGATTTCATACTGCTCTGCCAGCTCGGGATTGGATCATATCGCT 540
Db 481 CCTATGTCAGAGAGATTTCATACTGCTCTGCCAGCTCGGGATTGGATCATATCGCT 540

2Y 541 GAGAACATCTGTCTATACCTGGATGCCAAATCACTATGTGTGTGTAATTTGTGCAAG 600
Db 541 GAGAACATCTGTCTATACCTGGATGCCAAATCACTATGTGTGTGTAATTTGTGCAAG 600

2Y 601 GAATGTTACCGAGTACCTCTGATGCGATGCTGTGGAGAGCTTATCGAGAGATGCTC 660
Db 601 GAATGTTACCGAGTACCTCTGATGCGATGCTGTGGAGAGCTTATCGAGAGATGCTC 660

2Y 661 AGGACAGATTCTCTGTGGAGAGGCTTGGCAAGCAAGAGAGGATGGGACAGTATTTATTC 720
Db 661 AGGACAGATTCTCTGTGGAGAGGCTTGGCAAGCAAGAGAGGATGGGACAGTATTTATTC 720

2Y 721 AAAAAAACCCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT 780
Db 721 AAAAAAACCCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT 780

2Y 781 AAAAATTATCAAGACATTGAGACAAATAGAACTTAATTTGGAGATGTGGAAGACATAGTTTA 840
Db 781 AAAAATTATCAAGACATTGAGACAAATAGAACTTAATTTGGAGATGTGGAAGACATAGTTTA 840

```

```

841 CAGAGAAATTCATCTCCGAAGTGAAACAAAGAGGAGTTTACTGTTTACAGTATGATGAT 900
841 CAGAGAAATTCATCTCCGAAGTGAAACAAAGAGGAGTTTACTGTTTACAGTATGATGAT 900

901 CAGAAAATAGTAAGCGGCTTCGAGACAACACAAATCAAGATCTGGGATAAACACATTG 960
901 CAGAAAATAGTAAGCGGCTTCGAGACAACACAAATCAAGATCTGGGATAAACACATTG 960

961 GAATGCAAGCGAATTTCTACAGGCCATACAGGTTCTGCTCTCTGCTCCAGTATGATGAG 1020
961 GAATGCAAGCGAATTTCTACAGGCCATACAGGTTCTGCTCTCTGCTCCAGTATGATGAG 1020

1021 AGAGTGATCATAACAGGATCATCGGATTCACAGGTTCTGCTCTCTGCTCCAGTATGATGAG 1080
1021 AGAGTGATCATAACAGGATCATCGGATTCACAGGTTCTGCTCTCTGCTCCAGTATGATGAG 1080

1081 GAAATGCTTAAACACGTTGATTCAACATTTGGAAGAGTTCTGCACTTGGGTTTCAATAAT 1140
1081 GAAATGCTTAAACACGTTGATTCAACATTTGGAAGAGTTCTGCACTTGGGTTTCAATAAT 1140

1141 GGCATGATGTTGACCTCTCCAAAGATCGTTCCATTTGCTGATGGGATATGGCCCTCCCA 1200
1141 GGCATGATGTTGACCTCTCCAAAGATCGTTCCATTTGCTGATGGGATATGGCCCTCCCA 1200

1201 ACTGACATTTACCTCCCGAGGGTCTGTCGAGACACCGAGCTGCTCAATGTTAGAC 1260
1201 ACTGACATTTACCTCCCGAGGGTCTGTCGAGACACCGAGCTGCTCAATGTTAGAC 1260

1261 TTTGATGACAGTACATGTTCTGCAATCTGGGGATAGAACTATAAAGGTATGGAAACACA 1320
1261 TTTGATGACAGTACATGTTCTGCAATCTGGGGATAGAACTATAAAGGTATGGAAACACA 1320

1321 AGTACTTGTGAATTTGTAAGGACCTTAAATGGACACAAAGAGGCAATTCCTGTTTGCAG 1380
1321 AGTACTTGTGAATTTGTAAGGACCTTAAATGGACACAAAGAGGCAATTCCTGTTTGCAG 1380

1381 TACAGGACAGGCTGTGAGTGTGCTCATCTGACAACTATATCAGATATGGGACATA 1440
1381 TACAGGACAGGCTGTGAGTGTGCTCATCTGACAACTATATCAGATATGGGACATA 1440

1441 GAATGTGTGATCTTTACGAGTGTAGAGGCGCATGAGGAATTTGGTGGTGTATTCGA 1500
1441 GAATGTGTGATCTTTACGAGTGTAGAGGCGCATGAGGAATTTGGTGGTGTATTCGA 1500

1501 TTTGATTAACAGAGGATGATGAGTGGGCTTATGATGGGAAATAAAGTGGGATCTT 1560
1501 TTTGATTAACAGAGGATGATGAGTGGGCTTATGATGGGAAATAAAGTGGGATCTT 1560

1561 GTGGCTGCTTTGGACCCCGCTGCTGTCAGGAGCACTCTGCTACGGACCTTGTGGAG 1620
1561 GTGGCTGCTTTGGACCCCGCTGCTGTCAGGAGCACTCTGCTACGGACCTTGTGGAG 1620

1621 CATTCGGAGAGGTTTTCGACTACAGTTTGAATTTCCAGATTTGTCAGTATGATCAAT 1680
1621 CATTCGGAGAGGTTTTCGACTACAGTTTGAATTTCCAGATTTGTCAGTATGATCAAT 1680

1681 GATGACAAATCTTCATCTGGGCTTCTAAATGATCCAGTCCCAAGCTGAACCCCTCC 1740
1681 GATGACAAATCTTCATCTGGGCTTCTAAATGATCCAGTCCCAAGCTGAACCCCTCC 1740

1741 GCTTCCCTTCTCGAAACATACCTACATCTCCAGATAAATAACCATACACTGACCTCAT 1800
1741 GCTTCCCTTCTCGAAACATACCTACATCTCCAGATAAATAACCATACACTGACCTCAT 1800

1801 ACTTGGCCAGAGCCCATTAAGTTGGGATTTAAAGTATCTGCCAATACCAGATGAGC 1860
1801 ACTTGGCCAGAGCCCATTAAGTTGGGATTTAAAGTATCTGCCAATACCAGATGAGC 1860

1861 AACACAGTAAACAACTCAAACTACTGCCCAGTTTCCCTGGAGTAGCCGAGGAGGCTT 1920
1861 AACACAGTAAACAACTCAAACTACTGCCCAGTTTCCCTGGAGTAGCCGAGGAGGCTT 1920

1921 TGAGACTCTCTTGGGACACAGTTTGTCTGCACTCGGCCCGAGGAGGCTTCTATCAGACA 1980

```

```
Db      1921  TGAAGCTCTGTTGGGACACAGTTGCTGCGAGTGGGCCAGGCTCTACTCAGACA 1380
Qy      1981  ACTGACTGCTTCTAGTCTGCTATCAGAGATGCTCTTCTATCAATTTGTAATGATGGAA 2040
Db      1981  ACTGACTGCTTCTAGTCTGCTATCAGAGATGCTCTTCTATCAATTTGTAATGATGGAA 2040
Qy      2041  TTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db      2041  TTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Qy      2101  CAGACAAAGTGTCTATTAATAATATTAGTGTGTTTGGCAGAAAAA 2151
Db      2101  CAGACAAAGTGTCTATTAATAATATTAGTGTGTTTGGCAGAAAAA 2151

RESULT 3
US-10-038-010-7
; Sequence 7, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1710)
; OTHER INFORMATION: Human beta-TTCp1 : F-box containing protein with 7 WD40 repeats;
; OTHER INFORMATION: Part of SCF (Skp1/Cullin/F-box) complex E3 ubiquitin ligase; Imp
; OTHER INFORMATION: located in the degradation of b-catenin and IkBa
US-10-038-010-7

Query Match          79.5%; Score 1710; DB 14; Length 1710;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      70  ATGACCCCGCCGAGGCGGTGCTGCGAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
Db      1  ATGACCCCGCCGAGGCGGTGCTGCGAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 60
Qy      130  AGAGAGAGACTGTAATTAATGCGCAACCCCTAGGAAGATAATACCAAGAGAAATTCACCTT 189
Db      61  AGAGAGAGACTGTAATTAATGCGCAACCCCTAGGAAGATAATACCAAGAGAAATTCACCTT 120
Qy      190  AGACAGACATACACAGCTGTGCGACACTCTGTTAAACCAAGAAACAGTATGTTAGCA 249
Db      121  AGACAGACATACACAGCTGTGCGACACTCTGTTAAACCAAGAAACAGTATGTTAGCA 180
Qy      250  AGCACTGCTATGAAGACTGAGAAATTTGTGGCCAAAAAACAATTTGCCAATGGCACTTCC 309
Db      181  AGCACTGCTATGAAGACTGAGAAATTTGTGGCCAAAAAACAATTTGCCAATGGCACTTCC 240
Qy      310  AGTATGATTTGGCCAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAAGGAAGTCTG 369
Db      241  AGTATGATTTGGCCAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAAGGAAGTCTG 300
Qy      370  TGTGTCAAAATCTTTGAGCAGTGGTCTCAGAGTCAAGTCAAGTGAATTTGTGGAACATCTT 429
Db      301  TGTGTCAAAATCTTTGAGCAGTGGTCTCAGAGTCAAGTCAAGTGAATTTGTGGAACATCTT 360
Qy      430  ATATCCCAAATGTGTCATTACCAACATGGGCAATAACTCTGTATCTTAAACCTATGTTG 489
```

```
Db      361  ATATCCCAAATGTGTCATTACCAACATGGGCAATAACTCGTATCTTAAACCTATGTTG 420
Qy      430  CAGAGAGATTTTCAATACATGCTCTGCCAGCTCGGGATTTGGATCATATCTCGTGAACATT 549
Db      421  CAGAGAGATTTTCAATACATGCTCTGCCAGCTCGGGATTTGGATCATATCTCGTGAACATT 480
Qy      550  CTGTCACTACCTGGATGCAAAATCACTATGTGCTGCTGAACTTGTGTGCAAGAAATGGTAC 609
Db      481  CTGTCACTACCTGGATGCAAAATCACTATGTGCTGCTGAACTTGTGTGCAAGAAATGGTAC 540
Qy      610  CGAGTGACCTCTGATGGCATGCTGTGGAAGAACTTATCGAGAGAAATGTCTAGGACAGAT 669
Db      541  CGAGTGACCTCTGATGGCATGCTGTGGAAGAACTTATCGAGAGAAATGTCTAGGACAGAT 600
Qy      670  TCTCTGTGGAGAGGCTTGGCAGAACGAGAGATGGGACAGTATTATTCAAAAACAAA 729
Db      601  TCTCTGTGGAGAGGCTTGGCAGAACGAGAGATGGGACAGTATTATTCAAAAACAAA 660
Qy      730  CCTCTGACGGGAATGCTCTCTCCAACTCTTTTATAGACACTTTTATCCTAAAATTATA 789
Db      661  CCTCTGACGGGAATGCTCTCTCCAACTCTTTTATAGACACTTTTATCCTAAAATTATA 720
Qy      790  CAAGACATTGAGACAATAGAAATCTAATTGAGATGTGGAAGACATAGTTTACAGAAATT 849
Db      721  CAAGACATTGAGACAATAGAAATCTAATTGAGATGTGGAAGACATAGTTTACAGAAATT 780
Qy      850  CACTGCCGAAGTGAACAAGCAAGAGAGTTTACTGTTTACAGTATCATCATCAGAAAATA 909
Db      781  CACTGCCGAAGTGAACAAGCAAGAGAGTTTACTGTTTACAGTATCATCATCAGAAAATA 840
Qy      910  GTAAGCGGCTTCCGAGACAACACAACTCTGGATGGAATAAACAACATTTGGAATGCAAG 969
Db      841  GTAAGCGGCTTCCGAGACAACACAACTCTGGATGGAATAAACAACATTTGGAATGCAAG 900
Qy      970  CGAATTTCTACAGGCCATACAGGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1029
Db      901  CGAATTTCTACAGGCCATACAGGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Qy      1030  ATACAGGATCATCGGATTTCCAGGTCAGAGTGTGGGATGTAATACAGGTGAATGCTA 1089
Db      961  ATACAGGATCATCGGATTTCCAGGTCAGAGTGTGGGATGTAATACAGGTGAATGCTA 1020
Qy      1090  AACAGCTTGAATTCACCATTTGTGAAGCAGTTCTCTGACCTTGCCTTCAATATGSCATGATG 1149
Db      1021  AACAGCTTGAATTCACCATTTGTGAAGCAGTTCTCTGACCTTGCCTTCAATATGSCATGATG 1080
Qy      1150  GTGACCTGCTCCAAAGATCGTTCCATTTGCTGATGGATATGCCCTCCCACTGACATT 1209
Db      1081  GTGACCTGCTCCAAAGATCGTTCCATTTGCTGATGGATATGCCCTCCCACTGACATT 1140
Qy      1210  ACCCTCCGGAGGGTCTGCTGGGACACCGAGCTGCTGTCAATGTTGTAGACTTTGATGAC 1269
Db      1141  ACCCTCCGGAGGGTCTGCTGGGACACCGAGCTGCTGTCAATGTTGTAGACTTTGATGAC 1200
Qy      1270  AGTACATTTGTTCTGCTATCGGGATAGAACTATTAAGGTATGGACACAGTACTTGT 1329
Db      1201  AGTACATTTGTTCTGCTATCGGGATAGAACTATTAAGGTATGGACACAGTACTTGT 1260
Qy      1330  GAAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTTCAGTACAGGAC 1389
Db      1261  GAAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTTCAGTACAGGAC 1320
Qy      1390  AGGCTGTGAGTGTGCTCATCTGACACACTATCAGATTATGGGACATAGNATGTGT 1449
Db      1321  AGGCTGTGAGTGTGCTCATCTGACACACTATCAGATTATGGGACATAGNATGTGT 1380
Qy      1450  GCATGTTTACAGTGTTAGAAGGCCATGAGAAATTTGTCGTTGTATTTGATTTGATAAC 1509
Db      1381  GCATGTTTACAGTGTTAGAAGGCCATGAGAAATTTGTCGTTGTATTTGATTTGATAAC 1440
Qy      1510  AAGAGGATAGTCACTGGGGCTTATGATGGAAAAAATTAAGTGTGGGATCTTGTGGCTGCT 1569
Db      1441  AAGAGGATAGTCACTGGGGCTTATGATGGAAAAAATTAAGTGTGGGATCTTGTGGCTGCT 1500
```

2Y 1570 TTGGACCCCGGCTCCTGCGAGGACACTCTGTCTACGGACCCCTGTGGAGCATTCGGA 1629
Db 1501 TTGGACCCCGGCTCCTGCGAGGACACTCTGTCTACGGACCCCTGTGGAGCATTCGGA 1560
2Y 1630 AGAGTTTTCGACTACAGTTTGTATGAATTCAGATGTGTCAGTAGTTCACATGATGACACA 1689
Db 1561 AGAGTTTTCGACTACAGTTTGTATGAATTCAGATGTGTCAGTAGTTCACATGATGACACA 1620
2Y 1690 ATCCCATCTGGAGTTCCTAATGATCCAGTCCGCAAGCTGACCCCGGTTCCCT 1749
Db 1621 ATCCCATCTGGAGTTCCTAATGATCCAGTCCGCAAGCTGACCCCGGTTCCCT 1680
2Y 1750 TCTGAAACATACACCTACATCTCCAGATAA 1779
Db 1681 TCTGAAACATACACCTACATCTCCAGATAA 1710

RESULT 4
US-09-832-161-15
; Sequence 15, Application US/09832161
; Publication No. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercurio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinon
; APPLICANT: Davis, Matti
; APPLICANT: Hatzubai, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: NP-KB
; FILE REFERENCE: 860098.427
; CURRENT APPLICATION NUMBER: US/09/832.161
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/210,060
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 4230
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-161-15

Query Match 41.1%; Score 883.4; DB 12; Length 4230;
Best Local Similarity 74.8%; Pred. No. 5.2e-272;
Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;
QY 298 AATGGCAGTCCAGTATGATGTGCGCCAGCAAGCAAGCAACTCTCAGCAAGCTATGAAAG 357
Db 179 AATGGACATCATCTGTGATCTCTCCAGAAAGGCCATCAGAAAGCAATCAAAAAA 238
QY 358 GAAAGGAACCTGTGTCAAACTATTTGAGCAGTGGTGGAGTGGTGGAGTGGAGTGGAGTGG 417
Db 239 GAAAGGAACCTGTGTATTAATATTTTACAGTGGTCTGAAATCAGATCAAGTGGAAATTT 298
QY 418 GTGGACATCTTATATCCAAATGTCTATACCAATGAGGCAATAACTCGATCTT 477
Db 299 GTGGAAACATCTTATTTACGAATGTCTATATCAGCATGAGCATATTAATCTCTTACCTG 358
QY 478 AAACCTATCTTGCAGAGAGATTTCAATACTGTCTGCCAGCTCGGGATTTGGATCATATC 537
Db 359 AAGCCATGTGTCAGCGGAGCTTATTAACCGCTTTACCAAGCAAGGCTTAGATCAATA 418
QY 538 GCTGAGAACATCTGTCTACCTGATGCGCAAACTACATATGTGTGTGAACTGTGTGTC 597
Db 419 GAGAGAAACATCTTCTTGTGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 478
QY 598 AAGGATGTACCGAGTACCTCTGATGGCATGCTGTGGAGAGGCTTATCGAGAGATG 657
Db 479 AAGGATGTACCGAGTACCTCTGATGGCATGCTGTGGAGAGGCTTATCGAGAGATG 538

QY 658 GTACAGGACAGATTCTCTGTGAGAGGCTGCGAGAACGAAGAGATGGGACAGTATTTA 717
Db 539 GTACAGGACAGATTCTCTGTGAGAGGCTGCGAGAACGAAGAGATGGGACAGTATTTA 598
QY 718 TTCAAAAAACAAACCTCCCTGAGCGGAATGCTCCTCCCAACTCTTTTATAGAGACATTTAT 777
Db 599 TTTAAAAACAGACCCACAGATGGC-----CCTCCAAATTCATTTATAGGTCAATATAC 652
QY 778 CTTAAATTTATCAAGACATTTGAGACATTAAGATCTTAATTTGGAGATGTGAAACATAGT 837
Db 653 CCAAGATTTATCCAGGATATAGAGACTATAGAACTTAATCTAAGTGTCTACTCTTTACAGTAGAT 712
QY 838 TTACAGAGATTTCACTCCGGAAGTGAACAGCAAGCAAGAGTTTACTGTTTACAGTAGAT 897
Db 713 TTGAGAGATTTCACTCCGCTCTGAATAATGTAAGTGTCTACTCTTTACAGTAGAT 772
QY 898 GATCAAAAAATAGTAAGCGGCTTTGAGACAAACAATCAAGATCTGGGATAAAAAACACA 957
Db 773 GATCAAAAAATATCAGTGGCTACGAGATAATTTCTAATAGATATGGGATAAAACCCAGC 832
QY 958 TTGGAATCGAAGCAATTTCTCAGAGGCCATACAGTTTCACTCTCTCTCCAGTAGAT 1017
Db 833 CTGGAATGTTTGAAGTGTAAACAGGACACACAGGCTCTCTCTCTCTCTGCGAGTAGAT 892
QY 1018 GAGAGAGTATCATAACAGGATCATCGGATTTCCACGGTTCAGAGTGTGGGATGTAATAACA 1077
Db 893 GAGCGTGTCAATTGTAACCTGGCTCTTCAGATTTCTACGGTGAAGTGTGGGATGTAACACG 952
QY 1078 GGTGAATGCTAAACAGCTTGAATTCACGATGTGAAGCAGTTCTGCACTTGGCTTCAAT 1137
Db 953 GGTGAATGCTTAAACAGCTTGAATTCACGATGTGAAGCAGTTCTGCACTTGGCTTCAAT 1012
QY 1138 AATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
Db 1013 AATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072
QY 1198 CCAACTGACATTTACCTCCGAGGAGTGTGTCGGAACACCGAGTGTCTCAATGTTGTA 1257
Db 1073 GCGACCGACATCACTTTACCGCTGTCTCTGTTGGCCACCGGCTGCGCTCAATAGTA 1132
QY 1258 GACTTTGATGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
Db 1133 GACTTTGACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
QY 1318 ACAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
Db 1193 ACGAGCAGCTGTGAATTTGTTCTGATCTCTCAATGGGCAAGCGGGCAATGCTGTCTC 1252
QY 1378 CAGTACAGGACAGCGTGTGATGAGTGGCTCATCTGACACACTATCAGATATGGGAC 1437
Db 1253 CAGTACAGGATCGCTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312
QY 1438 ATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497
Db 1313 ATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372
QY 1498 CGATTTGATACAGAGAGTATGATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGG 1557
Db 1373 CGATTTGATACAGAGAGTATGATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGG 1432
QY 1558 CTGTGGCTGCTTTGGACCCCGCTCTCTGAGGAGCTCTGTCTACGAGCCCTTGTG 1617
Db 1433 TTGCAAGCTGCTCTTGACCTCGAGCCCGAGCAGCAGCATTTGTTTGGCAGCATTTGGTG 1492
QY 1618 GAGCATTCGGAAGAGTGTTCGACTACAGTTTGAATGAAATCCAGATGTTGAGTAGTCA 1677
Db 1493 GAAATTCGAGAGTGTTCGCTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552
QY 1678 CATGATGACAAATCCTCATCTGAGGCTCTCTAAATGATCCAGCTGCCAGCTGAACCC 1737
Db 1553 CATGATGACAAATCCTCATCTGAGGCTCTCTAAATGATCCAGCTGCCAGCTGAACCC 1612

610 CGAGTGACCTCTGATGSCATGCTGTGGGAGAGCTTATCGACAGAAATGGTCAGGACAGAT 669

494 GTACGCACTGATCCCTCTATGGAAGGACTTTCAGAAAGAGAGGGTGGGATCAGTACCTG 553
718 TTCAAAAAAACAACCTCTGACGGGATGCTCTCCCAACTCTTTTATAGAGCACTTTAT 777
554 TTTAAAAACAGACCCACAG-----ATGCCCTCCAAATTCATTTATAGTCAATTATAC 607
778 CCTAAATTTATACAGACATTGAGACAATAGAACTAATCGGATGTGGAAGACATAGT 837
608 CCAAGATTTATCCAGSATATAGAGACTATAGAACTAATCTGCGGTGTGGACGACAAAC 667
838 TTACAGAGAATTCACCTGCCGAGTGAACAAGCAAGAGGATTTACTGTTTACAGTATGAT 897
668 TTCCAGAGGATTCAGTGCCGCTCTGAAATAGTAGTAAGGTGTCTACTGTTTACAGTACGAT 727
898 GATCAGAAATAGTAGGCGCTTCGAGACACACAATCAAGATCTGSGATATAAAACACA 957
728 GATGAAAAAATTTATCAGTGGCC-TCGAGATAATTCATTAAGATATGGGATAAAACCGC 786
958 TTGGAATGCAAGGAATTCACAGGCCATACAGTTCAGTCCCTCTGTCCTCCAGTATGAT 1017
787 CTGGAATGTTGAAGTGTWACAGACACACARGCTCTGK-CTCTGTCGCGATATGAT 845
1018 GAGAGAGTATCATACAGGATCATCGGATTCAGGTCCAGGTGAGAGTGTGGATGTAAATACA 1077
846 GAKCGGTGATTTKTWACTGGCTCTTTCAGATCTACGNGAGAGTGTGGATGTGACACCG 905
1078 GGTGAATGCTAAACAGGTGATTCACCATTCGTGAGCAGTCTCTGC 1123
906 GTGAAAGNTCTTAACACATTTGATCCNCCNATTTGANGCTGATTTGC 951

RESULT 7
US-10-222-020-15
Sequence 15, Application US/10222020
Publication No. US20030175739A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ08C2
CURRENT APPLICATION NUMBER: US/10/222,020
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 10/116,016
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 09/764,848
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937

```
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08

Query Match      20.2%; Score 435.4; DB 12; Length 951;
Best Local Similarity 72.3%; Pred. No. 1.7e-128;
Matches 597; Conservative 5; Mismatches 216; Indels 8; Gaps 3;

QY 298 AATGCCACTTCCAGTATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATGAAAG 357
DB 134 AATGGAACATCATCTGTGATCGTCTCCAGAAAGAGGCCATCAGAGGAAACTATCAAAA 193
QY 358 GAAAAGGAACCTGTGTCAATACTTTGAGCAGTGTGTGAGGTCAGATCAAGTGAATTT 417
DB 194 GAAAAGACTTGTGATTAATAATTTTGACCAGTGTCTGAATCAGATCAAGTGAATTT 253
QY 418 GTGGAACATCTTATATCCAAATGTGTCAATTAACAACATGGGCACATAAACTCGTATCTT 477
DB 254 GTGGAACATCTTATTTCAGAATGTGTCAATTAACAACATGGGCACATAAACTCTTACCTG 313
QY 478 AAACCTATGTGAGAGAGAGATTTCAATACTGTCTGCCAGCTCGGGGATTTGATCATATC 537
DB 314 AAGCCCATGTTGACGCGGACTTTATTACCGCTTTACAGAGCAAGGCTTAGATCACATA 373
QY 538 GCTCAGAACATTTCTGTGATACCTGGATGCCAAATCACTATGTGTGCTGCTGAACTTTGTGC 597
DB 374 GCAGAAAAACATTTCTTGTGATGCGAGTCTCTGTGTGACGACAGCTGGTATGT 433
QY 598 AAGGAATGGTACCGAGTGACCTCTGATGGCATGCTCTGGAAGAGAGCTTATCGAGGAATG 657
DB 434 AAAGAATGGCAGGAGTGATCTCAGAAAGGAATGCTTTGGAAGAGAGCTGATTGAACGAATG 493
QY 658 CTCAGGACAGATTTCTGTGAGAGGCTTGGCAGAGGCTTGGCAGAGAGGATGGGACAGTATTTA 717
DB 494 GTACGCACTGATCCCTATGGAAGGACTTTGAAAGAGAGAGGCTGAGTACGTACCTG 553
QY 718 TTCAAAACAAACCTCTGTGAGGGAATGCTCTCCCAACTCTTTTATAGAGACCTTAT 777
DB 554 TTTAAAAACAGACCCACAG-----ATGGCCCTCCAAATTCATTTTATAGTCAATTATAC 607
QY 778 CCTAAAAATTATACAGACATTTGAGACAATAGAAATCTAATTTGGAGATGTGGAAGACATAGT 837
DB 608 CCAAGATTATCCAGGATATAGAGACTATAGATCTAATCTGGCGGTGTGAGACACACAC 667
QY 838 TTACAGAGATTTCACTCCGAAGTGAACAGCAAGAGGAGTTTACTGTTTACAGTATGAT 897
DB 668 TTGAGAGAGGATTCAGTCCGCTCTGAAAATAGTAAAGGTGTCTACTGTTTACAGTACGAT 727
QY 898 GATCAGAAAAATAGTAAAGCGGCTTCCAGAGACAAACAAATCAAGATCTGGGATAAAAACACA 957
DB 728 GATGAAAAAATTATCAGTGGCC--TCGAGATAATTCATTAAAGATATGGGATAAACCGC 786
QY 958 TTGAATGCAAGCGAATTTCTCAGAGCCATACAGGTTTCAAGTCTCTGTTCCAGTATGAT 1017
DB 787 CTGGAATGTTTGAAGTGTGTTWACAGGACACACARGCTCTGK-CTCTGTCTGAGTATGAT 845
QY 1018 GAGAGAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGTGGGATGTAAATACA 1077
DB 846 GAKCCGTGTTATTKTWTACTTGTGCTCTTCAAGATTTACCGGAGAGTGTGGGATGTGACACCG 905
QY 1078 GGTGAATGCTTAAACACAGTTGATTCACCATTTGGAAGCAGTTCTGC 1123
```

10b , 906 GTGAAAGNTCTTAACACATGATGATCCNCCNATTGGANGCTGATTGTC 951

RESULT 8

JS-10-116-016-15
 : Sequence 15, Application US/10116016
 : Publication No. US20030054379A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 : FILE REFERENCE: PTZ08C1
 : CURRENT APPLICATION NUMBER: US/10/116,016
 : CURRENT FILING DATE: 2002-04-05
 : Prior Application removed - See File Wrapper or Palm
 : NUMBER OF SEQ ID NOS: 53
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 15
 : LENGTH: 951
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc feature
 : LOCATION: (883)
 : NAME/KEY: misc feature
 : LOCATION: (913)
 : OTHER INFORMATION: n equals a,t,g, or c
 : NAME/KEY: misc feature
 : LOCATION: (931)
 : OTHER INFORMATION: n equals a,t,g, or c
 : NAME/KEY: misc feature
 : LOCATION: (935)
 : OTHER INFORMATION: n equals a,t,g, or c
 : NAME/KEY: misc feature
 : LOCATION: (941)
 : OTHER INFORMATION: n equals a,t,g, or c
 : US-10-116-016-15

Query Match 20.2%; Score 435.4; DB 14; Length 951;
 Best Local Similarity 72.3%; Pred. No. 1.7e-128;
 Matches 597; Conservative 5; Mismatches 216; Indels 8; Gaps 3;

QY	298	AATGCCACTTCAGTATGATTTGGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAAG	357
DB	134	AATGGAACATCTCTGTGATCTCTCCGAAAGAGGCCATCAGAAGCAACTATCAAAA	193
QY	358	GAAGAAGAACTGTGTCAAAATCTTTCAGCAGTGGTCAGAGTCAGATCAAGTGAATTT	417
DB	194	GAAGAAGACTTGTGATTAATTTTTCAGCAGTGGTCAGATCAAGTGAATTT	253
QY	418	GTGGAACATCTTATATCCAAATGTGTATACCAACTGGGCAATAACTGTATCTT	477
DB	254	GTGGAACATCTTATTTTCAGAAATGTGTATATCAGATGGACATATTAAGTCTTACCTG	313
QY	478	AAACCTATCTTCCAGAGAGATTTCAATCTGTCTGCCAGTCCGGGATTCGATCATATC	537
DB	314	AGCCCATGTTGACGGGACTTTTATACCCCTTTACCAAGCAGCTTAGATCAATA	373
QY	538	GTGGAACATCTTGTATCTGATCTGATGCAATCAATCAATCTGTGTCTGTAAGTGTGTC	597
DB	374	GCAGAAAACATCTTTCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG	433
QY	598	AGGATGTACCGAGTACCTCTGATGCGATCTGATGCGATCTGATGCGATCTGATGCGAATG	657
DB	434	AAAGATGCGAGGAGTGTATCTCAGAAGAAATGCTTTTGGAAAGAGCTGATTGAACGAATG	493
QY	658	GTGAGACAGATTTCTGTGAGAGAGCTTGGCAGAACCAAGAGATGGGACATATTA	717
DB	494	GTACGACTGATCCCTATGGAAGAGCTTTTTCAGAAAGAGGAGGATGGATCACTG	553
QY	718	TTCAAAAACAACTCTCTGACGGGAATGCTCTCCAACTCTTTTATAGAGCACTTTAT	777
DB	554	TTTAAAAACAGACCCACAG-----ATGCCCTCCAAATTCATTTATAGTCAATTATAC	607

RESULT 9

US-09-864-761-2778/c
 : Sequence 2778, Application US/09864761
 : Patent No. US20020048763A1
 : GENERAL INFORMATION:
 : APPLICANT: Penn, Sharron G.
 : APPLICANT: Rank, David R.
 : APPLICANT: Hanzel, David K.
 : APPLICANT: Chen, Wensheng
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 : FILE REFERENCE: Aeomica-X-1
 : CURRENT APPLICATION NUMBER: US/09/864,761
 : CURRENT FILING DATE: 2001-05-23
 : PRIOR APPLICATION NUMBER: US 60/180,312
 : PRIOR FILING DATE: 2000-02-04
 : PRIOR APPLICATION NUMBER: US 60/207,456
 : PRIOR FILING DATE: 2000-05-26
 : PRIOR APPLICATION NUMBER: US 09/632,366
 : PRIOR FILING DATE: 2000-08-03
 : PRIOR APPLICATION NUMBER: GB 24263.6
 : PRIOR FILING DATE: 2000-10-04
 : PRIOR APPLICATION NUMBER: US 60/236,359
 : PRIOR FILING DATE: 2000-09-27
 : PRIOR APPLICATION NUMBER: PCT/US01/00666
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00667
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00664
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00669
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00665
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00668
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00663
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00662
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00661
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: PCT/US01/00670
 : PRIOR FILING DATE: 2001-01-30
 : PRIOR APPLICATION NUMBER: US 60/234,687
 : PRIOR FILING DATE: 2000-09-21
 : PRIOR APPLICATION NUMBER: US 09/608,408

QY	778	CCTAAATTTATACAGACATTTGAGACAAATAGATCTAAATTTGGAGATGTGGAGACATAGT	837
DB	608	CCAAAGATTATCCAGGATATAGAGACTATAGAATCTAACTGCGGCTGTGGACGACACAC	667
QY	838	TTACAGAGAAATTCAGTCCGGAAGTGAACCAAGCAAGAGGATTTACTGTTTACAGTATGAT	897
DB	668	TTGAGAGGATTCAGTCCGCTCTGAAAATAGTAAAGGTGCTACTGTTTACAGTACGAT	727
QY	898	GATCAGAAAATAGTAAGCGCTTTTCAGAGCAACAAATCAAGATCTGGGATAAAACACA	957
DB	728	GATGAAAAAATATCAGTGGCC-TCGAGATAATCTATTAAAGATATGGGATAAAACGCG	786
QY	958	TTGGATCTCAAGCGAATTTCTCAGAGGCCATACAGTTTCACTCTCTCTCCAGTATGAT	1017
DB	787	CTGGAATGTTTGAAGTGTWACAGGACACACAGCTCTGK-CTCTGCTGCGCAGTATGAT	845
QY	1018	GAGAGAGTGATCATAACAGGATCATCGGATTCACCGTCAGAGTGTGGGATGTAATACA	1077
DB	846	GAKCGTGTATTTKTWACTGGCTCTTCAGATTTACGGNGAGAGTGTGGGATGTGACACCG	905
QY	1078	GGTGAATGCTAAACACAGTTCATTCAACCATTTGAGACAGTTCTGTC 1123	
DB	906	GTGAAAGNTCTTAACACATGATCCNCCNATTGGANGCTGATTGTC 951	

	Query Match	9.7%	Score 208.4	DB 9	Length 370
	Best Local Similarity	74.3% <td>Pred. No. 86-56 <td></td> <td></td> </td>	Pred. No. 86-56 <td></td> <td></td>		
	Matches 278	Conservative	0	Mismatches 9	Indels 5
					Gaps 1
QY	1027	ATCAATACACAGATCATCGGATTCACAGGTCAGAGTGGGATGTAATAACACAGTGAATG	1086		
Db	369	ATTGTAACCTGGCTCTTCGGATTCTACGGTGCAGAGTGTGGGAAGTGAACACAGGGTGAAGTT	310		
QY	1087	CTAAACACGTTTGATTCCACCAATTGTGAAGCAGGTTCTGCACATTTCGGTTTTCAATTAATGGCATG	1146		


```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011387.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.1
; OTHER INFORMATION: EST_HUMAN HIT: B1088274.1, EVALUATION 0.00e+00
; OTHER INFORMATION: NT HIT: G16306497, EVALUATION 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9UXB1, EVALUATION 5.00e-43
US-10-029-386-18709

Query Match
Best Local Similarity 7.6%; Score 162.6; DB 12; Length 250;
Matches 195; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

2Y 1059 AGTGTGGATGTAATACAGGTGAATGCTAAACAGGTGATTACCAATGTAAGCAGT 1118
Db 250 AGTGTGGATGTAAGACAGGTGAATGCTAAACAGGTGATTACCAATGTAAGCAGT 191
2Y 1119 TCTGCATTCGCTTCAATAAGGATGATGTCACCTGCTCCAAAGATGTTCCATTGC 1178
Db 190 ATTGCATTACGCTTCAAGCAATGATGATGATGATGATGATGATGATGATGATG 131
2Y 1179 TGTATGGATATGGCTCCCAACTGACATTAACCTCCGAGGATGCTGTCGACACCG 1238
Db 130 TGTGTGGATGATGGCTTCTGCGACCGACATCACTTTAGCGGCTGCTGTTGGCCACCG 71
2Y 1239 AGCTGTGCTCAATGTTGTAGATTTGATGACAGATGATGTTCTGATCTGGGATAG 1298
Db 70 GCTGTGCTCAATGTTGTAGATTTGATGACAGATGATGTTCTGATCTGGGATAG 11

2Y 1299 AACTATAA 1307
Db 10 GACCATCA 2

RESULT 14
US-09-918-995-15756
; Sequence 15756, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: FROM VARIOUS CDNA LIBRARIES
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15756
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(475)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-15756

Query Match
Best Local Similarity 7.3%; Score 157.4; DB 11; Length 475;
Matches 164; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

2Y 1966 GGTCTACTGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2025
Db 55 GGACACCCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 114
2Y 2026 GTGAATGATGGAACATTTAAACCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2085
```

```
Db 115 GNGAATGATGGAACATTTGGAAGTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 174
QY 2086 TTTTCCCATTTGGTTCCAGACAAAGGTGACTTATAAATAATATTAGTGTGTTGCCAGAA 2143
Db 175 TTTTCCCATTTGGTTCCAGACAAAGGTGACTTATAAATAATATTAGTGTGTTGCCAGAA 232

RESULT 15
US-10-245-618-13
; Sequence 13, Application US/10245618
; Publication No. US20030143582A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; APPLICANT: Strohmaier, Helmo
; APPLICANT: Spruck, Charles
; APPLICANT: Sangfelt, Olle
; TITLE OF INVENTION: HCD4 MODULATES CYCLIN E DEGRADATION
; FILE REFERENCE: TSRI 779.2
; CURRENT APPLICATION NUMBER: US/10/245,618
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/404,116
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/322,947
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-10-245-618-13

Query Match
Best Local Similarity 5.8%; Score 124.2; DB 12; Length 1620;
Matches 340; Conservative 0; Mismatches 328; Indels 9; Gaps 1;

QY 883 TGTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 942
Db 655 TGCTTACAGTTTGTGTTAAACCGAATAGTTAGTGTGTTCTGATGACAACTTTAAAGTT 714
QY 943 TGGATATAAACAACATTTGGAATGCAAGCGAATTTCTCACAGGCCATACAGGTTCCGTC 1002
Db 715 TGGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
QY 1003 TGTCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1062
Db 775 TCATCAAAATGAGAGACAAACATCATCATCATCATCATCATCATCATCATCATCATCAT 834
QY 1063 TGGATGTAATATACAGGTGAAATGCTAAACACGTTGATTCACCATTTGCAAGCAGTTCTG 1122
Db 835 TGGATGCAAGACTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
QY 1123 CACTTGGCTTTCAATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1182
Db 895 TGTATGATCTTCTCATGAAAAAAGAGTTGTTAGCGTTCTCGAGATGCCACTCTTTAGGTT 954
QY 1183 TGGATATGCGCTCCCACTGACATTTACCTCCGAGGTTGCTGTCGACACCGAGCT 1242
Db 955 TGGATATGAGACAGGCCAGTGT-----ACATGTTTGTGATGGTCACTGTTGCA 1005
QY 1243 GCTGTCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302
Db 1006 GCAGTCCGCTGTGTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
QY 1303 ATAAGTATGAGACACAGTACTTGTGATTTGTAGGACCTTAATGACACACAAACA 1362
Db 1066 GTAAAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
QY 1363 GGCATTGCTGTTTGCAGTACAGGACAGGCTGGTGTAGTGGTCTCATCTGACACACT 1422
```



```
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,621
; FILING DATE: 16-OCT-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: McKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: BCM-02999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; JS-08-951-621-2

Query Match 5.8%; Score 124.4; DB 4; Length 126;
Best Local Similarity 99.2%; Pred. No. 9.4e-31;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 511 CTGCAGCTCGGGATTGGATCATATCGTCAGAACATTCGTGCATACCTGGATGCCAAA 570
3b 1 CTGCAGCTCGGGATTGGATCATATTCGTGAGAACATTCGTGCATACCTGGATGCCAAA 60

2Y 571 TCACATGTGCTGCTGAACCTGTGTCAGGAATGGTACCGAGTGACCTGTGATGCGATG 630
3b 61 TCACATGTGCTGCTGAACCTGTGTCAGGAATGGTACCGAGTGACCTGTGATGCGATG 120
2Y 631 CTGTGG 636
3b 121 CTGTGG 126

RESULT 3
JS-08-914-999-7
; Sequence 7, Application US/08914999
; Patent No. 6346406
; GENERAL INFORMATION:
; APPLICANT: Rvazanov, Alexey G.
; APPLICANT: Hait, William N.
; APPLICANT: Pavuk, Karen S.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; TITLE OF INVENTION: AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/914,999
```

```
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 601-1-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5900
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Dictyostelium discoideum
; US-08-914-999-7

Query Match 3.8%; Score 82.2; DB 4; Length 2237;
Best Local Similarity 47.4%; Pred. No. 4.4e-16;
Matches 286; Conservative 0; Mismatches 308; Indels 9; Gaps 1;

QY 959 TGAATGCAAGCGAATTCACAGGCCATACAGGTCAGTCCCTCTGTCTCCAGTATGATG 1018
DB 1490 TGAATGTGTTCAACCTTGAAAGGTGATGAAGTCCAGTTCGAATCAATTTGTTATAATG 1549
QY 1019 AGAGATGATCAATACAGGATCATCGGATTCACGGTCAGAGTGGGATGTAATACAG 1078
DB 1550 ATCAATATTTGTTAGTGGTTCATCAGATCATCAATTAAGTTTGGGATTTAAAGAAAT 1609
QY 1079 GTGAATGCTAAACACAGCTTGATTCACCATTTGTGAAGCAGTTCCTGCACCTTGGCTT 1138
DB 1610 TAAGATGATTTTACTTTGGAGGTCATGATAAACCCTGCCATACGGTCTCTATTGAATG 1669
QY 1139 ATGCGATGATGGTGACCTGCTCCAAAGATCGTTCATTCGCTGTATGGGATATGCGCTCCC 1198
DB 1670 ATAAATATTTGTTAGTGGTTCCTCTGACAAACTATCAAAAGTTTGGGATTTGAAACTT 1729
QY 1199 CAATGACATTAACCTCCGGAGGTCCTGTCGGACACCGAGCTGCTGTCAATGTTGTAG 1258
DB 1730 TGAATGTAATATACCTTGAAAGTCATGCCAGA-----GCCGTCAAAACACTTT 1780
QY 1259 ACTTGTATGACAAGTACATGTTCTGCACTCTGGGATAGAACTATAAGGTATGGAACA 1318
DB 1781 GTATATCTGTGCAATATTTATTTAGTGGTTCAAATGATAAAACTATCAAGGTTTGGGAT 1840
QY 1319 CAAGTACTTGTGAATTTGTAAGGACCTTAAATGGACACAAAGAGGSCATTGCTCTTTGC 1378
DB 1841 TGAACACTTTTCGTTGTAACACTACACTCTAAAGGTCATATAATGGTCCACTATCT 1900
QY 1379 AGTACAGGACAGGCTGGTAGTGAGTGGCTCATCTGACACACTATCAGATTAAGGACA 1438
DB 1901 GTATATTAGTACCAATCTCAGAGTGGCTCTCTATGATAAAACTATAAGAGTTTGGAT 1960
QY 1439 TAGAATGTGTGCATGTTTACGAGTGTAGAAAGGCGCATGAGGAATTTGGTCCCTGTATTC 1498
DB 1961 TAAAGATTTAGATGTTCCGCTACTTTAAGAGGCCATGATAGATGGGTGGAACATGG 2020
QY 1499 GATTGTGATAACAGAGATGATGTCAGTGGGCGCTATGATGGGAAAAATTAAGTGTGGGATC 1558
DB 2021 TAATTTGTGATAAAATTTATTTACTGCTAGTACGATAAATAACAATTAATAATTTGGGAT 2080
QY 1559 TTG 1561
DB 2081 TAG 2083

RESULT 4
US-08-899-578-1
; Sequence 1, Application US/08899578
```

Patent No. 6087153

GENERAL INFORMATION:

APPLICANT: Greenwald, Iva
APPLICANT: Hubbard, E. Jane
TITLE OF INVENTION: SEL-10 AND USES THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1195 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,578
FILING DATE: 24-JUL-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 278-0525

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS

LOCATION: 91..1854

US-08-899-578-1

Query Match 3.1%; Score 67.2; DB 3; Length 2481;

Best Local Similarity 50.6%; Pred. No. 3.9e-11;

Matches 162; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

2Y 1273 TACATGTTTCTGCACTCTGGGATAGAACTATAAGGTATGGACACAAAGTACTTGTGAA 1332

Db 1018 TATATTGTTAGCGGTCCACTGATAGAACTGTAAAGTTTGGAGTACTGTAGATGGTTCA 1077

2Y 1333 TTGTGAAGCACTTAANTGGACACAAACGAGGCATTGCCTGTTTGCAGTACAGGACAGG 1392

Db 1078 CTTCTCTACACTTCAAGGACATACCTCCACTGTTGCATGATGGCTATGGCTGTGTTCC 1137

2Y 1393 CTGGTGTGAGTGGTCTCTTGACAACTATCAGATTATGGGACATAGAAATGTGTGCA 1452

Db 1138 ATACTTGTCACTGGATCAGAGATACCACTCTCGTGTATGGGACGTAGATCCGGACGT 1197

2Y 1453 TGTTACAGGTGTAGAGGCGCATGAGAAATGTGGTGTGTTTCAATTGATAACAG 1512

Db 1198 CACCTGGCAACTTACATGGCCCATCATGACCCGTTGCATGCGTTCAATTGATGAACA 1257

2Y 1513 AGGATGTGAGTGGGGCCCTATGATGGAAAAATTAAGTGTGGGATCTTGTGGCTGCTTTG 1572

Db 1258 ACTGTGTTTCGGGAGGATATGATTTTACCGTTTAAATTTGGAATGCTCATCTGGAGA 1317

2Y 1573 GACCCCGTGTCTCTGCGAGG 1592

Db 1318 TGTATCCGTACTCTGACCGG 1337

RESULT 5

US-08-283-917-8

, Sequence 8, Application US/08283917

Patent No. 5849557

GENERAL INFORMATION:

APPLICANT: ADACHI, HIDEKI
APPLICANT: TSUJIMOTO, MASAFUMI
APPLICANT: INOUE, KEIZO
APPLICANT: ARAI, HIROYUKI
TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
ADDRESSEE: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,917
FILING DATE: 03-AUG-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 209943/1993

FILING DATE: 03-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5849557man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2292-030-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 2085 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Bos taurus

FEATURE:

NAME/KEY: CDS

LOCATION: 844..2073

US-08-283-917-8

Query Match

2.4%; Score 50.8; DB 2; Length 2085;

Best Local Similarity 47.3%; Pred. No. 8.2e-06;

Matches 262; Conservative 0; Mismatches 277; Indels 15; Gaps 3;

QY 1001 TCTCTCCAGTATGATGAGAGTGATCATACAGATCATCGGATTCACGGTCAGAG 1060

Db 1184 TCATTTTCCATCTGTTTTCAGTGTTATGTTCTCTGCTTCAGAGGATGCTTACAAATTAAGG 1243

QY 1061 TGTGGGATGTAATACAGGTGAATGCTAAACACGTTGATTCCACCATTTGTAAGACAGTTTC 1120

Db 1244 TGTGGGATGATGAGACTGGAGATTTTGAACGAACTCTTAAGGGGCATACAGACTCTGTAC 1303

QY 1121 TGCATTCGGTTTCAATAATGG-----CATGATGGTGACCTGCTCCAAGATCGTTCCA 1174

Db 1304 AGGATATTTTCAATCGACACAGTGGCAAGCTTCTGGCTTCATGTTCTTCAGAGATATGACCA 1363

QY 1175 TTGCTGTATGGGATATGGCTCCCAACTGACATTAACCTCCGAGGGTCTGCTGGAC 1234

Db 1364 TTAAGCTATGGGATTTTCAGGGCTTTGAATGCTCAGAACCATTCATGGCCATGAC---C 1420

QY 1235 ACCGAGCTGCTGTCAATGTTGTAGACTTTTGATGACAAGTATGTTTCTGTCATCTGGG 1294

```

1421 ACAATGTTTCTTCAGTAGCCATCATGCCAATGAGATCATATAGTGTCTGCCTCAAGG 1480
1295 ATAGAACTATAAAGTATGGACACAGACTCTGTGAATTGTAAAGACCTTAAATGGAC 1354
1481 ATAAACATATAAAAAATGTGGGAAGTGAACACTGGCTACTGTGTGAAGACATTCACAGGAC 1540
1355 ACAAAAGGAGGCAATTCCTGTTTTCAGTACAGGGA-----CAGGCTGTTAGTGTGGCT 1408
1541 ACAGAGATGGTAGTATGTGTGGGCCCAATCAAGACGGCACTCTCATAGCAGCTGTT 1600
1409 CATCTGACAACTATCAGATPATGGGACATAGAAATGTGGTGCAATGTTTACGAGTGTAG 1468
1601 CCAATGACCAGACTGTGGCTGTATGGTGTGCTAGCAACAAGGAATGCAAGGCTGAGCTTC 1660
1469 AAGGCCATGAGGAATTTGGTGGCTTCTGATTCGATTGATACAGAGGATAGTCAGTGGGG 1528
1661 GAGAATGAGCATGTGGTAGNATGCAATTCCTGGGCTCCTGAAAGCTCATTTCTTCCA 1720
1529 CCTATGATGGAAAA 1542
1721 TCTCTGAAGCAACA 1734

```

RESULT 6

IS-08-961-716-8

Sequence 8, Application US/08961716

Patent No. 5880272

GENERAL INFORMATION:

APPLICANT: ADACHI, HIDEKI

APPLICANT: TSUJIMOTO, MASAFUMI

APPLICANT: INOUE, KEIZO

APPLICANT: ARAI, HIROYUKI

TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME

NUMBER OF INVENTION: AND GENE THEREOF

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &

ADDRESSEE: NEUSTADT, P.C.

STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.716

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/283,917

FILING DATE: 03-AUG-1994

APPLICATION NUMBER: JP 209943/1993

FILING DATE: 03-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5880272man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2292-030-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 2085 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

```

; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 844..2073
US-08-961-716-8

```

```

Query Match      2.4%; Score 50.8; DB 2; Length 2085;
Best Local Similarity 47.3%; Pred. No. 8.2e-06;
Matches 262; Conservative 0; Mismatches 277; Indels 15; Gaps 3;

```

```

QY 1001 TCTGTTCTCCAGTATGATGAGAGAGTATCATTAACAGGATCATCGGATTCACCGGTACAG 1060
DB 1184 TCATTTTCCATCCTGTTGTTTCAGTGTATGTTCTGCTTCAGAGGATGCTACAATTAAG 1243
QY 1061 TGTGGGATGAATACAGGTGAATGCTTAACACGTTGATTCCACCATGTTGAAGCAGTTC 1120
DB 1244 TGTGGGATGATGAGATGAGATTTTGAACGAACTTCTTAAGGGGCAATACAGACTCTGTAC 1303
QY 1121 TGCATTTGGCTTTCAATAATAG-----CATGATGGTGACCTGTCTCCAAAGATCGTTCCA 1174
DB 1304 AGGATATTTTCATTCGACCACAGTGGCAAGCTTCTGGCTTCATGTTCTGCAGATATGACCA 1363
QY 1175 TTGCTGTATGGGATATGGGCTCCCACTGACATTAACCTCCGGAGGCTGCTGGTGGAC 1234
DB 1364 TTAAGCTATGGGATTTTCAGGGCTTTGAATGATCAGAACCATGTCATGGCCATGAC---C 1420
QY 1235 ACCGAGCTGCTGTCAATGTTGTAGACTTTTGTATGACAAAGTACATTTGTTCTGCATCTGGG 1294
DB 1421 ACAATGTTTCTTCAGTAGCCATCATGCCCCAATGGAGATCATATAGTGTCTGCCCAAGG 1480
QY 1295 ATAGAACTATAAAGTATGGAACACAGTACTTGTGAATTTGTAGGACCTTAATGGAC 1354
DB 1481 ATAAACTATAAAATGTGGGAAGTGCAAACTGGTACTGTGTGAAGACATTCACAGGAC 1540
QY 1355 ACAAAAGGAGCATTCGCTGTTTGCAGTACAGGGA-----CAGGCTGGTGTAGTGGCT 1408
DB 1541 ACAGAGATGGGTACGTATGTTGGCGGCCCAATCAAGACGGCACTCTGATAGCCAGCTGTT 1600
QY 1409 CATCTGACAACTATCAGATTTATGGGACATAGAAATGTTGGTGCATGTTTACGAGTGTAG 1468
DB 1601 CCAATGACCAGACTGTGCGTGTATGGGTCTAGCAACAAGGAATGCAAGGCTGAGCTTC 1660
QY 1469 AAGGCCATGAGGAATTTGGTGGCTTGTATTCGATTTGTATAACAAGAGGATAGTCAAGTGGG 1528
DB 1661 GAGAACTAGCATGTGTTAGTATGCAATTTCTGGGCTCCTGAAAGCTCATATTTCTTCCA 1720
QY 1529 CCTATGATGGAAAA 1542
DB 1721 TCTCTGAAGCAACA 1734

```

RESULT 7

US-08-232-463-14/c

Sequence 14, Application US/08232463

Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEIFLINGER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

```
/
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703)836-9300
/ TELEFAX: (703)683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: pTZ9pt-P1s
/
/ US-08-232-463-14
/
/ Query Match          2.3%; Score 50; DB 1; Length 7218;
/ Best Local Similarity 4.9%; Pred.No. 3.5e-05;
/ Matches 20; Conservative 218; Mismatches 169; Indels 0; Gaps 0;
/
/ QY 586 GAACCTGTGTGCAAGAAATGTTACCGAGTACCTCTGATGCGCATGCTGTGGAAGAACTT 645
/   |||
/ Db 1448 GAAGAATTGGTACRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1389
/
/ QY 646 ATCGAGAGAAATGTCAGGACAGATCTCTGGAGAGCGCTGGCAGACAGAGAGGATGG 705
/
/ Db 1388 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1329
/
/ QY 706 GGACAGCATATTTTCAAAACAAACCTCTGACGGGAATGCTCTCCCAACTTTTAT 765
/   : : : : :
/ Db 1328 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1269
/
/ QY 766 AGAGCACTTTATCCCTAAATTAACAAGACATTTGAGACAAATAGAACTTAATTTGAGATG 825
/   : : : : :
/ Db 1268 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1209
/
/ QY 826 GGAAGACATAGTTTACAGAGAAATTCACCTGCCGAGTGAACAAAGAGGAGTTTACTGT 885
/   : : : : :
/ Db 1208 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1149
/
/ QY 886 TTACAGTATGATGACAGAAATAGTAAGCGGCTTCGAGACACACAAATCAAGATCTGG 945
/   : : : : :
/ Db 1148 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1089
/
/ QY 946 GATAAAACACATTTGAATGCAAGCGAATTCCTCAGAGCCCATACAG 991
/   : : : : :
/ Db 1088 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1043
/
/ RESULT 8
/ US-08-188-582-17
/ Sequence 17, Application US/08188582
/ Patent No. 553410
/ GENERAL INFORMATION:
/ APPLICANT: Tjian, Robert
/ APPLICANT: Comai, Lucio
/ APPLICANT: Dynlacht, Brian D.
/ APPLICANT: Hoey, Timothy
/ APPLICANT: Ruppert, Siegfried
/ APPLICANT: Tanese, Naoko
/ APPLICANT: Wang, Edith
/ APPLICANT: Weinzierl, Robert O.J.
/
/
/ TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
/ TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
/ STREET: 4 Embarcadero Center, Suite 3400
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-4187
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/188,582
/ FILING DATE: 28-JAN-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Osman, Richard A.
/ REGISTRATION NUMBER: 36,627
/ REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 781-1989
/ TELEFAX: (415) 398-3249
/ TELEX: 910 277299
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2152 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..2112
/
/ US-08-188-582-17
/
/ Query Match          2.2%; Score 47.8; DB 1; Length 2152;
/ Best Local Similarity 52.6%; Pred. No. 8e-05;
/ Matches 131; Conservative 0; Mismatches 112; Indels 6; Gaps 1;
/
/ QY 1265 ATGCAAGTACATTTGTTCTGTCATCTGGGATAGAACTATAAAGGTATGGAACACAACTA 1324
/   : : : : :
/ Db 1625 ATTCTAATTATTTGTTGCTACGGGCTCTGCAGACAGAACTGTGCGGCTCTGGGAGCTCTGA 1684
/
/ QY 1325 CTGTGAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTGCAGTA-- 1382
/   : : : : :
/ Db 1695 ATGTTAACTGTGTAAAGATCTTCACTGACACAAAGGGACCAATTCATCTCTTGACATTTT 1744
/
/ QY 1383 ----CAGGACAGGCTGGTAGTGCTCATCTGACACACACTATCAGATTATGGGACA 1438
/   : : : : :
/ Db 1745 CTCCTCAATGGGAGATTCCTGGCTACAGGACAAACAGATGGCAGAGTCTTCTTTGGGATA 1804
/
/ QY 1439 TAGAATGTTGTCATGTTTACGAGTGTAGAGGCAATTCAGGAAATGGTGGCGTTGTTTC 1498
/   : : : : :
/ Db 1805 TTGACATGTTTGTGATGTTGGAGAAATTAAGGCGCCACTGTATACAGTCTGTTCACTTA 1864
/
/ QY 1499 GATTGTATA 1507
/   |||
/ Db 1865 GGTTAGTA 1873
/
/ RESULT 9
/ US-08-646-715-17
/ Sequence 17, Application US/08646715
/ Patent No. 5637866
/ GENERAL INFORMATION:
/ APPLICANT: Tjian, Robert
/ APPLICANT: Comai, Lucio
/ APPLICANT: Dynlacht, Brian D.
/ APPLICANT: Hoey, Timothy
```

```

; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOEACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,715
; FILING DATE: 09-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/189,582
; FILING DATE: 28-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277239
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2152 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2112
; US-08-646-715-17

```

```

Query Match      2.2%; Score 47.8; DB 1; Length 2152;
Best Local Similarity 52.6%; Pred. No. 8e-05;
Matches 131; Conservative 0; Mismatches 112; Indels 6; Gaps 1;

QY 1265 ATGCAAGTACATTGTTCTGCACTGCGGATAGAACTATAAGTATGGAACAAGTA 1324
DB 1625 ATTCATATTATGTTGCTACGGCTCTGCACAGAACTGTGGGCTCTGGGACGCTCTGA 1684
QY 1325 CTTCGTAATTGTAAGGACCTTAAATGGACACAAACGAGGCAATCCCTGTTGCACTA-- 1382
DB 1685 ATGGTACTGTGAAGATCTTCACCTGGACACAGGACCAATTCACTTCCTTGACATTT 1744
QY 1383 ----CAGGACAGCTGGTAGTGAGTGGCTCATCTGACAACTATCAGATTATGGGACA 1438
DB 1745 CTCCTAATGGAGATTCCTGGCTACAGGAGCAACAGATGGCAGAGTCTCTTTGGGATA 1804
QY 1439 TAGAATCTGCTGATGTTTACGAGTCTTAGAAGGCCATGAGGAATTGGTGCTGTATTC 1498
DB 1805 TTGACATAGTTGATGTTGGAGAAATTAAAGGCCACACTGATACAGTCTGTTCACTTA 1864
QY 1499 GATTGTGATA 1507
DB 1865 GGTTTAGTA 1873

```

```

RESULT 10
US-08-232-463-14

```

```

; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
; US-08-232-463-14

```

```

Query Match      1.9%; Score 39.8; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 0.076;
Matches 14; Conservative 210; Mismatches 167; Indels 0; Gaps 0;

QY 1714 GATCCAGCTGCCAAGCTGAACCCCGCTCCCTTCTCGAACATACACCTACATCTCC 1773
DB 1052 GAGGAGCTTGGCATYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1111
QY 1774 AGATAATATACCATACACTGACCTCATCTGCCCCAGGACCATTAAGTTCGGTATT 1833
DB 1112 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1171
QY 1834 AACGTATCTGCCATACAGGATGAGCAACACAGTAACTAACTACCTCCAGTTT 1893
DB 1172 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1231
QY 1894 CCCTGGACTAGCGAGGAGCGGCTTGGAGACTCTGTGGACACAGTTCGTCTGAG 1953
DB 1232 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1291
QY 1954 TCGGCCGAGGACGGTCTTACTCAGCAACTGACTGCTCAGTCTGCTATCAGAACTGT 2013
DB 1292 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1351
QY 2014 CTTCTATCAATTGTGAATGATTGGAACCTCCCTCTCTCTCTCTCTCTCTCTCT 2073

```


; APPLICANT: Spaulding, Vikki
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 ; TITLE OF INVENTION: ENCODING THEM
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/738,367
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brown, Scott A.
 ; REGISTRATION NUMBER: 32,724
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8224
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 733 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-738-367-6

Query Match 1.8%; Score 38.6; DB 1; Length 733;
 Best Local Similarity 51.4%; Pred. No. 0.04;
 Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 2Y 1396 GTAGTGGTGGCTCTCAGACACTATCAGATTATGGGACATAGATGGTGGCATGT 1455
 Db 107 GTCACTAATCTTCTCAGGGAATATATGCTATTTTATCCATAGCATTTGGAAGCG 166
 2Y 1456 TTACGAGTGTAGAGCCATCAGGAATTGGTCGCTTATTCGATTGTGATCAAGAGG 1515
 Db 167 CAGGACACGTTAATGGGACATGATGCTGTTAGTAAGATCTGTGGCATGACACAGG 226
 2Y 1516 ATAGTCAGTGGGCTTATGAGAAATTAAGTGTGGATCTTGTGGCTGC 1568
 Db 227 CTATATCTGCATCGTGGGACTCTACAGTGAAGGTGCTGCTGCTGCTGCTGCTG 279

RESULT 14
 US-09-156-425-1
 ; Sequence 1, Application US/09156425B
 ; Patent No. 5962671
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Brenda F.
 ; APPLICANT: Cowsett, Lex M.
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF FAN EXPRESSION
 ; FILE REFERENCE: RTS-0009
 ; CURRENT APPLICATION NUMBER: US/09/156,425B
 ; CURRENT FILING DATE: 1998-09-18
 ; NUMBER OF SEQ ID NOS: 47
 ; SEQ ID NO 1
 ; LENGTH: 3380
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (23)..(2766)
 ; US-09-156-425-1

Query Match 1.8%; Score 38.6; DB 2; Length 3380;
 Best Local Similarity 51.4%; Pred. No. 0.11;
 Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 QY 1396 GTAGTGGTGGCTCTCAGACACTATCAGATTATGGGACATAGAAATGGTGGCATGT 1455
 Db 2071 GTCACTAATCTTCTCAGGGAATATATGCTATTTTATCCATAGCATTTGGAAGCG 2130
 QY 1456 TTACGAGTGTAGAGCCATCAGGAATTGGTCGCTTATTCGATTGTGATCAAGAGG 1515
 Db 2131 CAGGACACGTTAATGGGACATGATGCTGTTAGTAAGATCTGTGGCATGACACAGG 2190
 QY 1516 ATAGTCAGTGGGCTTATGAGAAATTAAGTGTGGATCTTGTGGCTGC 1568
 Db 2191 CTATATCTGCATCGTGGGACTCTACAGTGAAGGTGCTGCTGCTGCTGCTGCTG 2243

RESULT 15
 US-09-620-312D-1051
 ; Sequence 1051, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yungqing
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pt_FL_genes Version 1.0
 ; SEQ ID NO 1051
 ; LENGTH: 7518
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(7518)
 ; US-09-620-312D-1051

Query Match 1.7%; Score 36.8; DB 4; Length 7518;
 Best Local Similarity 53.5%; Pred. No. 0.75;
 Matches 77; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 QY 1242 TGCTGTCAATGTTGTAGACTTTGATGACAAAGTACATTTTCTGTCATCTGGGATAGAAC 1301
 Db 1563 TTCAGATATGGCAGTAAACTATGAGAATACAATGATGCTGGGGGAGCTGTGATAAAT 1622
 QY 1302 TATTAAGGTATGGAACACACAGTACTTGTGAATTTTAAGGACCTTAATGGACACACAGG 1361
 Db 1623 TATTAGAGTGTGGTGTGAGAACTTGTGCCAGTGTGCTGTGCTCCAGGACACACAGG 1682
 QY 1362 AGCAATTGCCTTTTGCAGTACAG 1385

db 1683 ATCATTACATCTTTACAGTTTAG 1706

Search completed: October 22, 2003, 21:59:58
Job time : 150 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2597	85.6	518	2	B48088	beta-transducin re
2	1635.5	53.9	701	2	T16607	hypothetical prote
3	690	22.7	506	2	T50211	WD-repeat protein
4	590.5	19.5	605	2	T38932	probable sulfur me
5	545	18.0	640	2	S49932	MET30 protein - ve
6	531.5	17.5	650	2	T46660	sulfur controller
7	520	17.1	579	2	T23703	hypothetical prote
8	513.5	17.1	267	2	S65507	hypothetical tri-p
9	455.5	15.0	1356	2	T18521	beta transducin-li
10	453	14.9	775	2	T45136	WD repeat protein
11	413.5	13.6	1227	2	AE1810	WD-40 repeat prote
12	399	13.2	779	2	S56245	cell division cont
13	396.5	13.1	703	2	T43557	F-box/WD-repeat pr
14	378.5	12.5	1189	2	AI2493	WD-repeat protein
15	375	12.4	1747	2	AC1842	WD-40 repeat prote
16	374	12.3	1526	2	AC3239	WD-40 repeat prote
17	373.5	12.3	1258	2	AI2155	WD-repeat protein
18	373.5	12.3	1683	2	AF2071	WD-40 repeat prote
19	361.5	11.9	677	2	AE1861	serine/threonine k
20	358.5	11.8	559	2	AB2202	hypothetical prote
21	354	11.7	409	2	S36113	lilS-1 protein - hu
22	354	11.7	410	2	S48052	platelet-activatin
23	341	11.2	515	2	T19487	hypothetical prote
24	339.5	11.2	777	2	T41075	hypothetical WD-re
25	337	11.1	1146	2	AS5532	myosin-heavy chain
26	336	11.1	676	2	AH2195	hypothetical prote
27	334	11.0	589	2	AG2400	WD-repeat protein
28	333.5	11.0	1711	2	AD1942	WD-40 repeat prote
29	332.5	11.0	317	2	T46032	WD-40 repeat regul

[illegible]

C;Function:
 A;Description: negatively regulates sulfur structural gene expression
 A;Note: scon-2+ expression is dependent on Cys3 function and the binding of Cys3 to the
 C;Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 17.5%; Score 531.5; DB 2; Length 650;
 Best Local Similarity 25.0%; Pred. No. 3.6e-13;
 Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;

QY 117 VEHLISOMCHYQGHCHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELV 176
 DB 99 LOGILSLQCFQPSFVSREVNEALKIDFISALPV-----ELAQVLCYLDTVSLTKAAQVS 154
 QY 177 KEWYRVTSQDMLKKKLTIERMVRTDSLWRGLAERRGWQYLFKNKP-----P 222
 DB 155 QRWRTLADSDAVWVRMCEQHVNK-----CTKCGWGLPLERKKLRWYTRQOLAKGP 208
 QY 223 DQNA-----PNSFYRA 234
 DB 209 QGRVTELADSHDSQDRSVNQHGKRPAAEAEEEDPIKKRCMAAAEASKAVTPKTRSMKA 268
 QY 235 LYPKLIQDIETIESNWRGCRHSQRIHCRSETSKGYVCLQYDDQKIVGSLRDNTIKIWDK 294
 DB 269 VY-----RDRQVSYNKNSEYKLSVL---KXGHEGVTCQLQDDNLIATGSDYTIKIWI 321
 QY 295 NTLCKRILTCHTGSVLCLOYDERVITGSSDSTVRVMDVNTGEMNLTIHHCEAVLHLR 354
 DB 322 ETECIRTVLGHTAG:RALQFDSDSKLISGLDHTIKWNWHTGECLSFAAHTDSVISVH 381
 QY 355 FNGMVMTCSDRSIAVMDWASPTDITLRRVLVGRRAAVV--VDFDDKIYVSAGDRTI 412
 DB 382 FDGHLLASGSDTKVIFDNSEKTYCLK-----GHSDWNVSTHVDIKSRTVFSASDDTTI 437
 QY 413 KVNTSTCEVRLNGHKGRI-----ACLVDRDLVS----- 445
 DB 438 KLMDLDRQIRTYEGHVGHVQVQLPPEYEPDEVLNGASQDNQDAMSVSQSGSPS 497
 QY 446 -----GSSD 449
 DB 498 MSHAQIERAGSPGSHSSHNLLPSSLPSGDEDVRHLYGSFAFVADSRPLPRPYMTSGLD 557
 QY 450 NTRLMDIEGACLVLEGHHEELVRCIRFDNKRIVSGAVDGKIKVMDLVAALDPRAPAGT 509
 DB 558 STWRLWDSATGRCRLTLFGHLEGWGLAGDTIRVIRSGANDGMVKTWE-----PRSG 608
 QY 510 LCLRTLVEHSGRVRFLQFQFQVSSSHDDTILWDF 546
 DB 609 KCDATYTGCGPVCVGLSDSLMASGSDGTIRLHSF 645

RESULT 7
 T22703
 hypothetical protein F55B12.3 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T22703
 R;Sims, M.
 submitted to the EMBL Data Library, September 1996
 A;Reference number: Z19602
 A;Accession: T22703
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-579 <WIL>
 A;Cross-references: EMBL:Z79757; PIDN: CAB02129.1; GSPDB: GN00023; CESP: F55B12.3
 A;Experimental source: clone F55B12
 C;Genetics:
 A;Gene: CESP: F55B12.3
 A;Map position: 5
 A;Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

Query Match 17.1%; Score 519.5; DB 2; Length 267;
 Best Local Similarity 38.0%; Pred. No. 8.8e-33;
 Matches 104; Conservative 58; Mismatches 83; Indels 29; Gaps 6;

QY 286 DNTIKIMDKNLECKRILTGTGTSVLCLOYDER--VITGSSDSTVRVMDVNTGEMNL 343
 DB 1 DRTVSVMDVNRFLYKLYGHSGSVLCDFCRNRLLVSGSSDSTIIIDWQNRPLKVY 60
 QY 344 IHHCEAVLHLRFNNGMVMVTCSDRSIAVW--DMASPTDITLRRVLVGHRAAVNVDFDK 401
 DB 61 FGHTDNLVGVVSENYIISSSRDHTARVRLDATSPAEACM-HVLRGHLASVNSVQYSK 119
 QY 402 --YIVSAGSDRTIKWNTSTCEFVRTLNGHKGRIACLOYDRDLVSGSSDNTIRLWDISC 459
 DB 120 TGLIVTASSDRTLRTWDTITGHGICIRIIHAHQGIACQYNGKFIIVSGSSDLTIRIFASS 179

QY 72 KTKLANGTSSMI-----VPQQRK--LSAYEKEKEL-----CVKYFEQNSQDQVE 115
 DB 27 ESSYNGSSSYNADKLSSSRPLQHKLDLSASPSRNDLNPVRVHILIAFKDLSSAEQMD 86
 QY 116 FVEHLISOMCHYQGHCHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELV 175
 DB 87 AFTRLQBSNNNTIQLRAIEHPHFQDFLSCLPV-----ELGKILHLNLTGVDLLKVAQV 142
 QY 176 KEWYRVTSQDMLKKKL-IERMV-----RTDSLWRGLAERRGWQYLFKNKPPDGNAP 227
 DB 143 SKNWKLISEIDKIWKSILGVBEFKHPDPDTRVTGAMQGTAAIAG-----VTIFDHIQF 195
 QY 228 PN-SFYRALYKLIQDI-----ETESNWRGCRHSQRIHCRSETSKGYVCL 273
 DB 196 CDLVNVRHLKLOKFGDIFERAADKSRVLRADKTEKWNANPIMGSV-VRGHEDHVTCK 254
 QY 274 QYDDQKIVSGURDNTIKIWDKNTLECKRILTGTGTSVLCLOYDE--RVITGSSDSTVRV 331
 DB 255 QIHDDVLVTGSDNTLKWICDKGEVMTLVGTGVTGTSQISQCGRYIVSGSTORTVRV 314
 QY 332 WDVNTGEMNLTIHHCEAVLHLRFNNGMVMVTCSDRSIAVMDWASPTDITLRRVLVGHRA 391
 DB 315 WSTDVGSLLHLTQGHSTVRCMAGSILVTGSDTTLRVWDVSEGRHLA---TLGHHRA 371
 QY 392 AVNVVDPDKIIVSAGDRTIKWNTSTCEFVRTLNGHKGRIACLOYRDR--LVVSGSSD 449
 DB 372 AVRCQFQDGTTVVSGGYDFTVKIWNHTGRCIRTLTGHNNRVYSLFESERSIVCSGSLD 431
 QY 450 NTRLMDI---ECGACLVLEGHHEELVRCIRFDNKRIVSGAYDGKIKVMDLVAALDPRAP 506
 DB 432 TSIRVWDITRPEGQECVALLQGHSTLTSQWQLRGNILVSCNADSHVRVWDI-----H 483
 QY 507 ACTLCRLTVLHSGRVRFLQ-FDEQFQIVSSSHDDTILWID 545
 DB 484 EGT-CVHMLSGHRAITSLQWFGNVMVATSSDDGTVKLWD 522

RESULT 8
 S62507
 hypothetical trp-asp repeat-containing protein - fission yeast (Schizosaccharomyces pombe
 C;Species: Schizosaccharomyces pombe
 C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C;Accession: T38502; S62507
 R;Jones, L.; Murphy, L.; McNeil, A.; Simpson, I.; Harris, D.; Barrell, B.G.; Rajandream,
 submitted to the EMBL Data Library, October 1995
 A;Reference number: Z21798
 A;Accession: T38502
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-267 <JO2>
 A;Cross-references: EMBL:Z66525; NID:G1044926; PIDN:CAA91423.1; PID:G1044927; GSPDB:GN00
 C;Genetics:
 A;Map position: 1
 A;Superfamily: unassigned WD repeat proteins; WD repeat homology
 P;18-51/Domain: WD repeat homology <WD1>
 F;104-137/Domain: WD repeat homology <WD2>

Query Match 17.1%; Score 519.5; DB 2; Length 267;
 Best Local Similarity 38.0%; Pred. No. 8.8e-33;
 Matches 104; Conservative 58; Mismatches 83; Indels 29; Gaps 6;

QY 286 DNTIKIMDKNLECKRILTGTGTSVLCLOYDER--VITGSSDSTVRVMDVNTGEMNL 343
 DB 1 DRTVSVMDVNRFLYKLYGHSGSVLCDFCRNRLLVSGSSDSTIIIDWQNRPLKVY 60
 QY 344 IHHCEAVLHLRFNNGMVMVTCSDRSIAVW--DMASPTDITLRRVLVGHRAAVNVDFDK 401
 DB 61 FGHTDNLVGVVSENYIISSSRDHTARVRLDATSPAEACM-HVLRGHLASVNSVQYSK 119
 QY 402 --YIVSAGSDRTIKWNTSTCEFVRTLNGHKGRIACLOYDRDLVSGSSDNTIRLWDISC 459
 DB 120 TGLIVTASSDRTLRTWDTITGHGICIRIIHAHQGIACQYNGKFIIVSGSSDLTIRIFASS 179

QY 460 GACLRVLEGEHEVLCIRPFNKIVSGAYDGKIKWDLVAALDPRAPAGTCLRLTLVBEH- 518
 Db 180 GKLLRLMQGHEDLIRTVFNEFIVSGYDGTGVIRWN-----FNTGEQHC 224
 QY 519 -----SGRYERLQFDEFQIVSSSHDITLINDF 546
 Db 225 VLHNSRNRVFGQFDRRIIACHSSEILVWNF 258

RESULT 9
 T18521
 beta transducin-like protein - Podospora anserina
 C:Species: Podospora anserina
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18521
 R:Saupé, S.; Turcq, B.; Begueret, J.
 Gene 162, 135-139, 1995
 A:Title: A gene responsible for vegetative incompatibility in the fungus Podospora anserina
 A:Reference number: Z18944; MUID:96009891; PMID:7557402
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1356 <SAU>
 A:Cross-references: EMBL:L28125; NID:g607002; PID:g607003; PIDN:AAA85775.1
 C:Genetics:
 A:Gene: het-el
 A:Introns: 761/3

Query Match 15.08; Score 455.5; DB 2; Length 1356;
 Best Local Similarity 31.74; Pred. No. 8.2e-27;
 Matches 121; Conservative 64; Mismatches 131; Indels 66; Gaps 15;

QY 226 APPNSFYRLY----PKIIQDIETIESNRCGRHSL-----QRHCRS- 264
 Db 802 APTDSMIKKIKKEPGWISTISVEAEWNACTQTLEHGSSVLSVAFSADQORVASGD 861
 QY 265 -----ETSKG-----VYCLQY--DDQKIVSGLRDNTIKIWDNLTCEKRL 303
 Db 862 DTKIKINDTASGTQTLEHGSSVMSVAFSPDRERVASGSDDKTIKIWDASGCTOTL 921
 QY 304 TGTGTVLCLQY--DERVIITGGSDSVRVWVNTGEMLNTLIHCEAVLHLRFN-NGMM 360
 Db 922 EHGGRVQSVAFSPDQORVASGSDDKTIKIWDASGCTQTLEHGSSVLSVAFSPDQOR 981
 QY 361 VTC-SKDRSTAVMDASPTDITLRVLVGHRAVNVVDF--DDKIYVSASGRTIKWNT 417
 Db 982 VASGSGDKTIKINDTASG--TCTQLEHGSSVMSVAFSPDQORVASGSDDKTIKINDT 1038
 QY 418 STCEVFTLNGHKGIACLQYR--DRLVVSGSDNTIRLWDIECGACLRVLEGEHEVLCR 475
 Db 1039 ASGTCTQLEHGSSVMSVAFSPDQORVASGSDDKTIKIWDASGCTQTLEHGSSVMS 1098
 QY 476 IRF--DNKRIVSGAYDGKIKWDLVAALDPRAPAGTCLRLTLVHSGRVRIQF--DEFQ 531
 Db 1099 VAFSPDQORVASGSDITKIND-----AASGT-CTQLEHGSSVMSVAFSPDQOR 1149
 QY 532 IVSSSHDITLINDFLNDPAAQ 553
 Db 1150 VASGSDITKINDAASGCTQ 1171

RESULT 10
 T45136
 WD repeat protein popl [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C:Accession: T45136; T40157
 R:Kominami, K.; Toda, T.
 submitted to the EMBL Data Library, September 1996
 A:Description: Fission yeast WD repeat protein Popl is involved in maintenance of ploidy
 A:Reference number: 222925
 A:Accession: T45136

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-775 <KOM>
 A:Cross-references: EMBL:Y08391; PIDN:CAA69671.1
 A:Experimental source: strain h- 972
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.
 submitted to the EMBL Data Library, March 1998
 A:Reference number: Z21842
 A:Accession: T40157
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-775 <WOO>
 A:Cross-references: EMBL:AL022103; PIDN:CAA17898.1; GSPDB:GN00067; SPDB:SPEC2G2.18
 A:Experimental source: strain 972h; cosmid c2G2
 C:Genetics: SPEC2G2.18
 A:Gene: SPCC2G2.18
 A:Map position: 2
 A:Note: popl+

Query Match 14.9%; Score 453; DB 2; Length 775;
 Best Local Similarity 25.8%; Pred. No. 5.8e-27;
 Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;

QY 9 QEKALKFMNSSEBDC-----NNGEPPIKIPEKNSLRQTY-NSCARLCLNQET 56
 Db 121 REKCLKRNSLSLNLHANKRFLFNSQSDGKNQNETPEFTNYSNVFYPNNCDKSEVASET 180
 QY 57 VC-----LASTAMKTCNVAKTKLANGTSSMIVPKORKLSA----- 92
 Db 181 TSLDAPNNSVNYSVFSPNLLGNDKTKRQSPFSSSSSHNSLHEPVIYDFSSNPISHP 240
 QY 93 ----SYEKEKEL----CVKYFEQMSSEDOVFVHLSIQMCHYQHGHINSYKPMLODF 144
 Db 241 SNHLSQKNAVLKLAQLISSFEKLPESVRQYLLFELLSRGKHAQVNIHKILLPIFOKNF 300
 QY 145 ITALPARGLDHTAENILSYLDAKSLCAELACKWYR-VTSDGMLWKKL-----I 193
 Db 301 LTGFPA----ETNLVLIHLDAPSLCAVSQSHHWIKLVSSNEELWKSFLKOGFFWDSI 356
 QY 194 ERMVRTDLSMRGLAERRGQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETIESNRCG 253
 Db 357 DSKIRTMCLEQLSA-----CAIMKRVYFRHFNLRW--- 389
 QY 254 RSLQRIHCRSETSKGYVC---LOYDDOKIVSGLRDNTIKIWDNLTCEKRLTGHGTVS 310
 Db 390 LHAPEKIKCSPPIHGRLITKIQDDDKIIVSTCSPRINTYDTKTGVLKSLSEHEGTV 449
 QY 311 LCLQYDERVITIGSSDSTVRVWVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSDKRSIA 370
 Db 450 WTEYVVDITLVGTSDTRTVRWDLTGECKQVFGYGTSTIRCIVQGNQSTTDTD--- 505
 QY 371 VMDMASPTDITLRVLVGHRAVNVVDFDDKVIYSASGRTIKWVN----- 416
 Db 506 -----DVEKE-----NRPASNDANSMPPIYISSSRDCTIRLWSLPCLDPPFPVNVNE 552
 QY 417 -----TSTCEVFTLNGHK---RGIACQYRDLRVVSGSSONTIRLWDIECGAC 462
 Db 553 NPQDNDFTSATNPPYIRTLRGHTDSVREACL---GGLIVSASYDGLTAVMKASTGVC 609
 QY 463 LRVLGEHEBELVRCIRFDNKR--IVSGAYDGKIKWDLVAALDPRAPAGTCLRLTLVHSG 520
 Db 610 LHLVLRGHVGRVSVTINFPSRQOCISAGTDAKIRIWNL-----ESGEL-LQTLHGHSN 660
 QY 521 RVPRLOQDEFQIVSSS--HDDTILIWD 545
 Db 661 LVSQVTFNQNILVASAPDTSIRVWD 687

RESULT 11
 AE1810
 WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

QY	336	IGEMWILIIHCEAVLHLRFNNGMIVISKOKSIAVWNASFIDILKRLVVGCHRAAVN	395
Db	452	KG-----CCT-----HVFEGHNSTVRC	468
QY	396	VFDD---KYIVSASGDRTIKWANT-----STCC-----FVRLTNG	428
Db	469	LDIVEYKNKIYIVTSGRDNTLHVWKLPRSSVPDHCEBHDYPLVFHTPEENPYFVGVLRG	528
QY	429	HKGIACIQRLRLVVGSSDNTIRLWDIECGACRLVEGHBELVRCIRPDN--KXIVSG	486
Db	529	HNASVETVSGHNIYVSGSYDNTLIWVDYAQMKCIYILSGHTDRITYSTIYDHERKCIISA	588
QY	487	AYDGKIKWDL-----VAALOPRAPAGTL--CLRTLVEHSGRVFLRQDEFQIVSS	535

Db 589 SMDTIRWLDENIMNNGECVATNSAPCAKILGAMVTLQHTALVGLRLSDKFLYSA 648
QY 536 SHDDTILWDLNDPAQAQAPRSPRTVY 566
Db 649 AADGSRGWD-AND-----YRKFSY 668

RESULT 13
F43557
N:Alternate names: proteolysis factor sudlp
C:Species: Schizosaccharomyces pombe
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
A:Accession: T43557; T38794; T43798
R:Wolf, D.A.; Jackson, P.K.
A:Description: Fission yeast pop2 encodes a novel F-box/WD-repeat protein involved in th
A:Reference number: 222576
A:Accession: T43557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-703 <WOL>
A:Cross-references: EMBL:AF038867; PIDN:AAB95480.1
R:Gentiles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1996
A:Reference number: 221812
A:Accession: T38794
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-703 <GEN>
A:Cross-references: EMBL:Z98602; PIDN:CAB11275.1; GSPDB:GN00066; SPDB:SPAC4D7.03
A:Experimental source: strain 972h-; cosmid C4D7
R:Jallepalli, P.V.; Tien, B.; Kelly, T.J.
Proc. Natl. Acad. Sci. U.S.A. 95, 8159-8164, 1998
A:Title: Sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Runt proteins fo
A:Reference number: 222686; MUID:98318628; PMID:9653157
A:Accession: T43798
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-703 <JAL>
A:Cross-references: EMBL:AF064515; NID:G3293382; PIDN:AAAC39496.1; PID:G3293383
C:Genetics:
A:Gene: pop2; sud1; SPAC4D7.03
A:Map position: 1
C:Function:
A:Description: required to prevent spontaneous re-replication

Query Match 13.1%; Score 396.5; DB 2; Length 703;
Best Local Similarity 23.3%; Pred. No. 1.2e-22;
Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

QY 14 KFMSSRED--CNGGEPPEK-----IPEKSLRQTYNSCARLCL-----NOETVCL 59
Db 96 KFNESENVSVCLOHSPDPSPVPGFDSILVPGQNF-LYSHSLPPLKIISIDRNNRIKLD 154

QY 60 ASTAMKTENCVAKTKLANGTSSMIVPQKRLSASYEK-EKELCVKYFPGWSESDQVEFVE 118
Db 155 NSISSNSDNPFPSPKV--DTSNTVSPGSPISDELDLNLQSIQVOTFEDLPBGIO--SYAF 211

QY 119 HLISQMHYQGHG-INSYKPMQLQRFITAPARGLDHIAENILSYLDAKSLCAAEVLCK 177
Db 212 FQLRSNCNRMILLNECEPLLKXILNLFP-----SIVQILLNLDHLSFLSCLVSP 267

QY 178 EWRVYV---TSDGMLWKKLIERMVRTDSLWRGLAERGGQVLFKNKPPDGNAPP-----N 229
Db 268 TWRNILDVHTS---YWKHMF-----SLFGQINENDW-----KYANENLNRPFLHND 312

QY 230 SFYRSLYPKLIQDITETTESNRCGRHSLOP-----RHLNKRKWLFPSPSHLSPPHVPNFMITSLLLH 360
Db 313 QISDDYFEIFK-----RHLNKRKWLFPSPSHLSPPHVPNFMITSLLLH 360

QY 260 -----IHRSETSK-GVYCLQYDDQXIVSGLRDNTIKIWKNT 296
Db 361 KDRITITSGSGTIOHNAITGVLEARLEHGKGVAVKIHENTLVSGSIDKTVRVVNIK 420

QY 297 LCKRILTGHGTSGVLCQY-----DERVITGSSDSTVRVMDVNTGEML 340
Db 421 AKCTHIFRGLHSIRCLLEILVPSRLIRHGVEIVPEPDQYIVSGSRDHTLRV----- 472

QY 341 NTLIHCEAVLHFRNNGMMVTCSDKRSIAVMDWASPTDITLRRVLVGHRAAVNVVDFD 400
Db 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNEFYFV---HTLIGHTDSVRTISGV 517

QY 401 KYIVSASGDRITIKVWNTSTCEFFVTLNGHKRGACLOYRD--RLVVGSSDNTIRLWDIE 458
Db 518 DVLVSGSDSSIRIWRVSTGECYLHGRHSRLIYSLVLEPERNICSGSMKIRVWDLS 577

QY 459 CGACLRVLEGHEELVRCIRFDNKRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEH 518
Db 578 TCTCKVYLEGHDAPFTLLNVFQNELISSADSTIRWD---LNTGKP---LMVLPSN 628

QY 519 SGRVRLQDFEFQIVSSSHDDTILWD 545
Db 629 SCYISFVSDHKII-SGNDGSKLWD 654

RESULT 14
AI2493
WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AI2493
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2493
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1189 <KUR>
A:Cross-references: GB:BA000020; PIDN:BA078213.1; PID:gl7135667; GSPDB:GN00180
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr7129
A:Genome: plasmid

Query Match 12.5%; Score 378.5; DB 2; Length 1189;
Best Local Similarity 26.7%; Pred. No. 6.4e-21;
Matches 96; Conservative 63; Mismatches 107; Indels 93; Gaps 10;

QY 280 IVSGLRDNTIKIWKNTLECKRILTGHGTSGVLCQYD--ERVITGSSDSTVRVMDVNTG 337
Db 745 LAGSFDQNVKLMIDHTGKVMTLQGHGTGVVTSVAFNPKDNLGSGYDQSVKVRDKTG 804

QY 338 EMLNTLIHCEAVLHFRN----- 356

Db 805 RCLDTLKXHTNRINVSVAHFHQCHLFSGGDDHAAKIHELGTGQCIKTFQGHSNATYIAH 864

QY 357 ---NGMMVTCSDKRSIAVMD--MASPTDITLR-----RLVVGHRAAVNVVDFD--KYIVS 405

Db 865 NWEHSLLAGSHEDQTIKLDNLNLSPHKSNVNTHPFRILQHSNRVFSVFSSTGQLLAS 924

QY 406 ASGDTIKVWNTSTCEFFVTLNGHKRGACLOYR--DLVVGSSDNTIRLWDIECCACL 463
Db 925 GSADRTIKLWSPHTQCCLHTLHGHSWVAFAFSLDDKLLASGYDHTVKIWDVSSQCL 984

QY 464 RVLGHEELVRCIRF--DNKRIVSGAYDGKIKVMDL-----VAALDP 503
Db 985 QTLQHGPGSLAVASCQKTLFSSGYEKLQKQWDETGYCLOQWEADSNRVAVSRD 1044

QY 504 RAPAGT-----LCRLTVEHSGRVRLQF--DEFQIVSSSHDDTILWD 545

Db 1045 NQVLTGDDSVVRLMDIGKGVCTRFSGHTSQVICILFTKDCRRMISSSSDRIKIWN 1103

RESULT 15

AC1842
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AC1842
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC1842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1747 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA077807.1; PID:G17135261; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0283

Query Match	12.4%	Score 375;	DB 2;	Length 1747;	
Best Local Similarity	32.7%	Pred. No. 2.le-20;	Mismatches 107;	Indels 16;	Gaps 7;
Matches	84;	Conservative 50;			
Qy	276	DDQKIVSGLRDNTIKIWDKNTLECKRILTHTGTVSLCLOY--DERVIITGSSDSTVRVWD	333		
Db	1487	DGKTIATASADNTIKLWDSOTQOLIKTLTGKDRITTLFSFHPDNTIAGSGADKTIKIWR	1546		
Qy	334	VNTGEMLNTLIHCEAVLHLRF--NNGMMVTCSDRSIAVMDASPTDITLRRVLVGHRA	391		
Db	1547	VNDGQLRLTGTGNDVTSVNFSPDQFLASGSTNTVKIWQ----TDGRLIKNTIGHGL	1602		
Qy	392	AVNVVDF--DDKXIVSAGDRTIKVWNTSTCEFVRTLNHGKRGIAQLQYR--DRLVWSGS	447		
Db	1603	AJASVKFSPDSHTLASASWDNTIKLQVTDGKLNKLNHIDGVTLSLSPSPDGEILASGS	1662		
Qy	448	SDNTIRLWDIECGACLVLEGHEELVRCIRF--DNKRIVSGAYDGKIKVWDLVAALDPRA	505		
Db	1663	ADNTIKLWNLPNATLLKTLHGHPGKINTLAFSPDGKTLSSGGEDAGVMVWNL--DLDDL	1720		
Qy	506	PAGTLCRLTIVHSGRV	522		
Db	1721	QQGCDRITDYLQHSNV	1737		

Search completed: October 22, 2003, 10:26:03
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

1M protein - protein search, using sw model

run on: October 22, 2003, 10:07:11 ; Search time 26 Seconds

(without alignments)

1029.161 Million cell updates/sec

Title: US-09-601-168B-2

Perfect score: 3034

Sequence: 1 MDPAAVLQKALKFMSSE.....PAAQAEPRSPSRYYTISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3006	99.1	605	FW1A_HUMAN	Q9V297 homo sapien
2	2597	85.6	518	TRCE_XENLA	Q91854 xenopus lae
3	2384.5	78.6	542	FW1B_HUMAN	Q9UKB1 homo sapien
4	1638.5	54.0	665	LI23_CAEEL	Q09990 caenorhabdi
5	690	22.7	506	POFB_SCHPO	Q09855 schizosacch
6	590.5	19.5	605	POF1_SCHPO	P87053 schizosacch
7	575	19.0	678	SCOB_ENEHI	Q00659 emericella
8	545	18.0	640	MT30_YEAST	P39014 saccharomyc
9	531.5	17.5	650	SCO2_NEUCR	Q01277 neurospora
10	520	17.1	579	SE10_CAEEL	Q91794 caenorhabdi
11	477.5	15.7	684	CDG2_CANAL	P53699 candida alb
12	455.5	15.0	1356	HET1_PODAN	Q00808 podospora a
13	453	14.9	775	POPI_SCHPO	P87060 schizosacch
14	399	13.2	779	CDG4_YEAST	P07834 saccharomyc
15	396.5	13.1	703	POP2_SCHPO	O14170 schizosacch
16	392	12.9	732	KMHB_DICDI	P90648 dictyosteli
17	374	12.3	1526	YY46_ANASP	Q8YR11 anabaena sp
18	373.5	12.3	1258	YS00_ANASP	Q8YTC2 anabaena sp
19	373.5	12.3	1683	YL24_ANASP	Q8YV57 anabaena sp
20	354	11.7	409	LISI_HUMAN	P43034 homo sapien
21	354	11.7	409	LISI_MOUSE	P43035 mus musculu
22	353	11.6	409	LISI_BOVIN	P43033 bos taurus
23	341	11.2	515	YK24_YEAST	P25382 saccharomyc
24	337	11.1	1146	KMWA_DICDI	P42527 dictyosteli
25	334.5	11.0	422	FW2_HUMAN	Q9UKT8 homo sapien
26	325.5	10.7	361	WDS_DROME	Q9V3J8 drosophila
27	324.5	10.7	334	WDR3_HUMAN	Q9UGP9 homo sapien
28	318	10.5	422	FW2_MOUSE	P60584 mus musculu
29	313.5	10.3	376	YKY4_CAEEL	Q17963 caenorhabdi
30	312.5	10.3	714	YJL2_YEAST	P47025 saccharomyc
31	307.5	10.1	742	PKWA_THECU	P49695 thermomonos
32	307.5	10.1	1693	Y163_SYNY3	Q55563 synechocyst
33	306	10.1	704	T2D4_DROME	P49846 drosophila

34	298.5	9.8	800	1	T2D4_HUMAN	Q15542 homo sapien
35	294.5	9.7	1249	1	APAF_RAT	Q9EPV5 rattus norv
36	292	9.6	659	1	YK16_YEAST	P36130 saccharomyc
37	290.5	9.6	614	1	TU11_SCHPO	Q09715 schizosacch
38	289	9.5	1249	1	APAF_MOUSE	O88679 mus musculu
39	287.5	9.5	327	1	GBUP_BRANA	Q39336 brassica na
40	287	9.5	1248	1	APAF_HUMAN	O14727 homo sapien
41	286.5	9.4	327	1	GBUF_ARATH	O24456 arabidopsis
42	285.5	9.4	798	1	T2D4_YEAST	P38129 saccharomyc
43	283	9.3	473	1	PRP5_SCHPO	O13615 schizosacch
44	283	9.3	713	1	TUPI_YEAST	P16649 saccharomyc
45	281.5	9.3	682	1	TUPI_KUJLA	P56094 kluyveromyc

ALIGNMENTS

RESULT 1

ID	FW1A_HUMAN	STANDARD;	PRT;	605 AA.
AC	Q9V297; Q9V213;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	F-box/WD-repeat protein 1A (F-box and WD-repeats protein beta-TrCP)			
DE	(E3RS1KappaB) (pikappaBalpha-E3 receptor subunit).			
GN	BTRC OR FBXW1A OR FBXW1A OR BTRCP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=99075339; PubMed=9859996;			
RA	Varon A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,			
RA	Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;			
RT	"Identification of the receptor component of the IkappaBalpha-			
RT	ubiquitin ligase."			
RL	Nature 396:559-564 (1998).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Lymphoid;			
RX	MEDLINE=98323370; PubMed=9660940;			
RA	Margottin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,			
RA	Thomas D., Strebel K., Benarous R.;			
RT	"A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu			
RT	connects CD4 to the ER degradation pathway through an F-box motif."			
RL	Mol. Cell 11:565-574 (1998).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RX	MEDLINE=20003060; PubMed=10531035;			
RA	Cenciarelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,			
RA	Pagano M.;			
RT	"Identification of a family of human F-box proteins."			
RL	Curr. Biol. 9:1177-1179 (1999).			
RN	[4]			
RP	CHARACTERIZATION.			
RX	MEDLINE=99145464; PubMed=9990852;			
RA	Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Ellledge S.J.,			
RA	Harber J.W.,			
RT	"The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically			
RT	with phosphorylated destruction motifs in I-kappa-B-alpha and			
RT	beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro."			
RL	Genes Dev. 13:1270-1283 (1999).			
RN	[5]			
RP	INTERACTION WITH PHOSPHORYLATED CTNNB1.			
RX	MEDLINE=22072105; PubMed=12077367;			
RA	Sadot E., Conacci-Sorelli M., Zhurinsky J., Shnizer D., Lando Z.,			
RA	Zharhary D., Kam Z., Ben-Ze'ev A., Geiger B.;			
RT	"Regulation of S33/S37 phosphorylated beta-catenin in normal and			
RT	transformed cells."			
RL	J. Cell. Sci. 115:2771-2780 (2002).			
CC	!- FUNCTION: Substrate-recognition component of the SCF (SKP1-CUL1-F-			

box protein) ubiquitin ligase complex, which mediates the ubiquitination of proteins involved in cell cycle progression, signal transduction and transcription. Regulates the stability of CTNNB1 and participates in Wnt signaling.

-!- PATHWAY: Ubiquitin conjugation; third step.

-!- SUBUNIT: Interacts directly with SKP1 in the SCF complex. Interacts specifically with phosphorylated CTNNB1 and NFKB1A, ubiquitination substrates.

-!- SUBCELLULAR LOCATION: Cytoplasmic.

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=2;

Name=1;

isoId=Q9Y297-1; Sequences=Displayed;

Name=2;

isoId=Q9Y297-2; Sequences=VSP 006764;

-!- SIMILARITY: Contains 1 F-box domain.

-!- SIMILARITY: Contains 7 WD repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF101784; AAD08702.1; -.

EMBL; Y14153; CAA74572.1; -.

EMBL; AF129530; AAF04464.1; -.

Genew; HGNC:1144; BTRC.

MIM; 603482; -.

GO; GO:0004840; F-ubiquitin conjugating enzyme activity; TAS.

GO; GO:0007165; P-signal transduction; TAS.

GO; GO:0006511; P-ubiquitin-dependent protein catabolism; TAS.

InterPro; IPR001810; F-box.

InterPro; IPR001680; WD40.

Pfam; PF00646; F-box; 1.

Pfam; PF00400; WD40; 7.

PRINTS; PR00320; GPROTEINBRPT.

ProDom; PD000018; WD40; 4.

SMART; SM00256; FBOX; 1.

SMART; SM00320; WD40; 7.

PROSITE; PS00181; FBOX; 1.

PROSITE; PS00678; WD_REPEATS_1; 6.

PROSITE; PS00082; WD_REPEATS_2; 7.

PROSITE; PS00294; WD_REPEATS_REGION; 1.

UniProt; UniProtKB; P12345; Ubiquitin conjugation pathway; Repeat; WD repeat; Alternative splicing.

DOMAIN; 190 228 F-box.

REPEAT; 301 338 WD 1.

REPEAT; 341 378 WD 2.

REPEAT; 381 418 WD 3.

REPEAT; 424 461 WD 4.

REPEAT; 464 503 WD 5.

REPEAT; 505 541 WD 6.

REPEAT; 553 590 WD 7.

VARSPLIC; 17 52 Missing (in isoform 2).

FT; FTID=VSP_006764.

SEQUENCE 605 AA; 6886 MW; 4C67F3B7E400FD37 CRC64;

Query Match 99.1%; Score 3006; DB 1; Length 605;

Best Local Similarity 94.0%; Pred. No. 8 6e-220;

Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

2y 1 MDPAEVAVLEKALKFM-----NSSEREDC 24

3b 1 MDPAEVAVLEKALKFMPSRLWGLCSLADSPSLRCLYNPGTGAITAFONSSEREDC 60

2y 25 NNGEPRKPIPEKNSLRQYNQSCARCLNQETVCLASTAMKTNCKVAKTKLANGTSSMIV 84

3b 61 NNGEPRKPIPEKNSLRQYNQSCARCLNQETVCLASTAMKTNCKVAKTKLANGTSSMIV 120

2y 85 PKQKLSASYEKEKELCVKFEQWSESDQVEFVEHLISQMHYQGHINSYLPKPLQDF 144

Db 121 PKQKLSASYEKEKELCVKFEQWSESDQVEFVEHLISQMHYQGHINSYLPKPLQDF 180

Qy 145 ITALPARGLHIAENILSYLDAKSLCAELVCKEWYRTSDGMLKKGLIERVMTDSLWR 204

Db 181 ITALPARGLHIAENILSYLDAKSLCAELVCKEWYRTSDGMLKKGLIERVMTDSLWR 240

Qy 205 GLAERRGQGYLFPKNKPPDGNAPNSFYRALYPKIQTDIETTESNRCGRHSLQIHCRS 264

Db 241 GLAERRGQGYLFPKNKPPDGNAPNSFYRALYPKIQTDIETTESNRCGRHSLQIHCRS 300

Qy 265 ETSKGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTHTGSLVCLQYDERVITGS 324

Db 301 ETSKGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTHTGSLVCLQYDERVITGS 360

Qy 325 SDSIVRVMDVNTGEMNTLIHCEAVLHFRNNGMVTCKDRIAVWDMASPTDITLRR 384

Db 361 SDSIVRVMDVNTGEMNTLIHCEAVLHFRNNGMVTCKDRIAVWDMASPTDITLRR 420

Qy 385 VLVGHRAAVNVDFDDKIVVSASGDRTIKVMNTSTCEFVRTLNHGRGIACIQYRDLVV 444

Db 421 VLVGHRAAVNVDFDDKIVVSASGDRTIKVMNTSTCEFVRTLNHGRGIACIQYRDLVV 480

Qy 445 SGSSDNTIRLWDIECGACLRVLEGHELVRCIFENKRIVSGAYDGIKWMDLVAALDPR 504

Db 481 SGSSDNTIRLWDIECGACLRVLEGHELVRCIFENKRIVSGAYDGIKWMDLVAALDPR 540

Qy 505 APAGTLCRLTVEHSGRVFRLQDFEIVSSSHDDTILWDFLNDPAAQAEPPSPRTY 564

Db 541 APAGTLCRLTVEHSGRVFRLQDFEIVSSSHDDTILWDFLNDPAAQAEPPSPRTY 600

Qy 565 TYISR 569

Db 601 TYISR 605

RESULT 2

TRCB XENLA STANDARD; PRT; 518 AA.

ID TRCB XENLA STANDARD; PRT; 518 AA.

AC Q91854; P70037; P70038;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Beta-TrCP (Beta-transducin repeat-containing protein).

GN FBXW1 OR BTRCP.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

OC NCBI_taxid=8355;

OX [1]

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=9330289; PubMed=8393141;

RA Spevak W., Keiper B.D., Strata C., Castanon M.J.;

RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase are rescued by Xenopus CDNA encoding N-ras or a protein with beta-transducin repeats.";

RT with beta-transducin repeats.";

RL Mol. Cell. Biol. 13:4953-4966(1993).

RN [2]

RP SEQUENCE OF 302-518 FROM N.A.

RX MEDLINE=97109804; PubMed=8952061;

RA Hudson J.W., Alarcon V.B., Elinson R.P.;

RT "Identification of new localized RNAs in the Xenopus oocyte by differential display PCR.";

RL Dev. Genet. 19:190-198(1996).

CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated proteins and promotes their ubiquitination and degradation. May participate in Wnt signaling.

CC -!- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX (BY SIMILARITY).

CC -!- MATURATION STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO


```

JR PRINTS; PRO0320; GPROTEINBRPT.
JR ProDom; PD000018; WD40; 4.
JR SMART; SM00256; FBOX; 1.
JR SMART; SM00320; WD40; 7.
JR PROSITE; PS00181; FBOX; 1.
JR PROSITE; PS00678; WD REPEATS 1; 5.
JR PROSITE; PS00682; WD REPEATS 2; 7.
JR PROSITE; PS00694; WD REPEATS REGION; 1.
CW Developmental protein; Cell cycle; Cell division;
CW ubl conjugation pathway; Repeat; WD repeat.
JT DOMAIN 81 127
JT REPEAT 220 257 WD 1.
JT REPEAT 260 299 WD 2.
JT REPEAT 301 337 WD 3.
JT REPEAT 343 380 WD 4.
JT REPEAT 383 420 WD 5.
JT REPEAT 423 460 WD 6.
JT REPEAT 472 509 WD 7.
JT MUTAGEN 441 441 G->R; IN LIN-32 (RH293).
JT SEQUENCE 665 AA; 75916 MW; BF3E9AF5F12ECC CRC64;

Query Match 54.0%; Score 1638.5; DB 1; Length 665;
Best Local Similarity 63.7%; Pred. No. 2.6e-116;
Matches 316; Conservative 59; Mismatches 92; Indels 29; Gaps 3;

2Y 88 RKLASAYEKEKELCVKPEOMSEDOVEFVHLISOMCHYOGHINSYKLPMLQDFITA 147
2b 27 KFLSDYLOCHGLEHLEVLKSEHQDPKIVHRLSHYOLGKVDNFIRPMLQDFISN 86
2Y 148 LPARGLDHIAENILSYLDAKSLCAELAVCKEYRVTSQDMLWKXLIERMVTSLSMGLA 207
2b 87 LPA---HLVELILFNVNSLSKCEEVSTWRCALARGQHWKLIKBNVRSDSLWGLS 142
2Y 208 ERKGWQYLV-----FNKPPDGNAPSPNSFVALYKPIQDIETIE 247
2b 143 EKRWQDFLNISRDMSVRIKCFNYDNIKRDQLDQILMHVYKLYPKIRDIHND 202
2Y 248 SNMRCGRHSLQRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIKDNTLBECKRLTGT 307
2b 203 NNWKGNYKWTINQSENSKGVYCLQYDDQKIVSGLRDNTIKIKDNTLBECKRLTGT 262
2Y 308 GSVLCLOYDERVLIIGSSDSYRVWDVNTGEMNTLIHCEAVLHLPNNGMVTCSDR 367
2b 263 GSVLCLOYDNVRIIGSSDSYRVWDVNTGEMNTLIHCEAVLHLPNNGMVTCSDR 322
2Y 368 STAVWDMASFTDITRRVLVGHRAAVNVVDFDKYIVSASGDRITKYVNTSTCFEVRTLN 427
2b 323 STAVWDMVSPRDTITRRVLVGHRAAVNVVDFDKYIVSASGDRITKYVNTSTCFEVRTLN 382
2Y 428 GHRGTACLOYDRDLVVSQSSDNTLWDEGACLRVLEGHELVRCIRFDKRIYVSGA 487
2b 383 GHRGTACLOYDRDLVVSQSSDNTLWDEGACLRVLEGHELVRCIRFDKRIYVSGA 442
2Y 488 YDGKIKVMDLVAALDPAPAGTLCRLTVAHSGRVFRQLQDFEFOIVSSSHDDTLIWDFL 547
2b 443 YDGKIKVMDLVAALDPAPAGTLCRLTVAHSGRVFRQLQDFEFOIVSSSHDDTLIWDFL 502
2Y 548 NDPAAQAEPRPSRT 563
2b 503 DAP-----PSGLPSST 513

RESULT 5
ID POB_SCHPO STANDARD; PRT; 506 AA.
ID Q09855; Q9P7V1;
ID 01-FEB-1996 (Rel. 33, Created)
ID 16-OCT-2001 (Rel. 40, Last sequence update)
ID 28-FEB-2003 (Rel. 41, Last annotation update)
ID F-box/WD-repeat protein pof11.
ID POF11 OR SPAC23E6.01 OR SPAC30.05.
ID Schizosaccharomyces pombe (fission yeast).

```

Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomycetes.
 NCBI_TaxID=4896;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21668955; PubMed=11809834;
 RA Katayama S., Kitamura K., Lehman A., Nikaido O., Toda T.;
 "Fission yeast F-box protein Pof3 is required for genome integrity and
 telomere function.";
 RL Mol. Biol. Cell 13:211-224 (2002).
 [2]
 SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 Skellern J., Simmonds M., Squares R., Squares S., Stevens K.,
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Woodward J., Vackaert G., Aert R., Robben J., Grymonprez B.,
 Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 "The genome sequence of Schizosaccharomyces pombe.";
 Nature 415:871-980 (2002).
 RL
 CC -!- SIMILARITY: Contains 1 F-box domain.
 CC -!- SIMILARITY: Contains 7 WD repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AB061694; BAB55543.1; -;
 EMBL; AL136538; CAB66464.1; -;
 EMBL; Z66525; CAA91423.1; -;
 PIR; T50211; T50211
 GenDB_SPombe; SPAC23E6.01; -;
 InterPro; IPR001810; F-box.
 InterPro; IPR001680; WD40.
 Pfam; PF00646; F-box; 1.
 Pfam; PF00400; WD40; 7.
 PRINTS; PRO0320; GPROTEINBRPT.
 ProDom; PD000018; WD40; 1.
 SMART; SM00256; FBOX; 1.
 SMART; SM00320; WD40; 8.
 PROSITE; PS00181; FBOX; 1.
 PROSITE; PS00678; WD REPEATS 1; 3.
 PROSITE; PS00682; WD REPEATS 2; 6.
 PROSITE; PS0294; WD REPEATS_REGION; 1.
 Repeat; WD repeat.
 DOMAIN 70 116 F-BOX.
 REPEAT 219 256 WD 1.

```

27 REPEAT      259      298      WD 2.
28 REPEAT      301      338      WD 3.
29 REPEAT      345      386      WD 4.
30 REPEAT      388      426      WD 5.
31 REPEAT      427      464      WD 6.
32 REPEAT      468      505      WD 7.
33 SEQUENCE      506 AA; 58257 MW; CEF34DAEFBC2E10 CRC64;

Query Match
Best Local Similarity 30.4%; Score 60; DB 1; Length 506;
Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

2Y 67 ENCVAK-----TKLANGTSMIVPKOKKLSASYEKEKELCVKFEQWSSRSDOVERFVEHLIS 122
2b 8 KNVSVKSDLTSCDSFSTSSPVCNPLS-----HENNRIDLRDLA 50

2Y 123 QMCHYQGHINSYKMLQORDFITALPARGLHIAENILSYLDAKSLCAELCKEYWRV 182
2b 51 SLSKEGVAVYNHVRSLFTDFTEVEP-----EVSRLVFSYLDQLDCKCKLMSKRWKRL 106

2Y 183 TSDGMLWKLI-----ERMVRTDSLWRG-----LAERRGWG----- 213
2b 107 LEDEGIKALYMKQKGFVNENVLNBEFEMWRTHKFQPFENFLKQNMIGYGTWMLPQ 166

2Y 214 QYLFQKPPGPNAPPNSFYRALPKIIOQIETIESNRCGRHSLORIHCRS----- 264
2b 167 QFIF-----DSNGRPLNWSYLY-----KEHAHLDNWRHGRFLVSTFNPNPFIAPPQDF 217

2Y 265 -EYSGKGYCQYDQKIVSGRLNTIKIWDKNTLECKRILTGTGSLVLCLOYDER--VII 321
2b 218 RATLDSYVCQYDEIMVSGSKDRTVSVMDVNSRFLYKLYGHSGSVLCLCFRRNLLV 277

2Y 322 TGSSDSTVRVWNTVEMLNTLHHCBAVLHLPFNMGMMVTCSDKRSIAVW--DMASTD 379
2b 278 SGSSDSTIIIMQNRPLKVPYFGHTDNLGVVSENYIISSRDHTARVRLDATSPA 337

2Y 380 ITRRLVLVGHRAAVNVVDFDK--YIVSASGDRTIKWNTSCFVRTLNGHKRGIAQLQ 437
2b 338 ACM-HVLRGHLASVNSVSSKGLIVTASSDRTLTWDTITGHGIRIHAHQGIACAQ 396

2Y 438 YRRLVVGSSDNTIIRLWIEGACLRVLEGHLELVRCPFNKRVISGAVDGKIKWDL 497
2b 397 YNGKFIIVSGSDLTIRIFEASSGKLRLMLQGHEDIRTVRENDKIVSGGYDGTVRWN- 455

2Y 498 VAALDPRAPAGTCLRLTVEH-----SGRVRLQDFQFVSSSHDDTILWDF 546
2b 456 -----FNTGECHVLHNSRNSRVFGLQFDHRIIACHTSHSEILVWNF 497

RESULT 6
ID POFL SCHPO STANDARD; PRT; 605 AA.
AC P87053;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pofl (Skp1-binding protein 1).
GN POFL OR SBP1 OR SPAC57A10.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
DC Schizosaccharomycetales; Schizosaccharomycetaceae;
DC Schizosaccharomycetes.
CX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RA Katayama S., Kitamura K., Lehmann A., Nikaido O., Toda T.;
RT "Fission yeast F-box protein Pofl is required for genome integrity and
telomere function."
RL Mol. Biol. Cell 13:211-224 (2002).
RP SEQUENCE FROM N.A.

```

```

RC MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Pearson D., Quay M.A., Rabinowitsch E.,
RA Oliver K., O'Neil S., Munhall K., Murphy L., Niblett D., Odell C.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Art R., Robben J., Grynoprez B.,
RA Weljens I., Vantreels E., Kieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Leleure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880 (2002).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -!- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 F-box domain.
CC -!- SIMILARITY: Contains 7 WD repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB032410; BAA84528.1; --
DR EMBL; Z94864; CAB08168.1; --
DR PIR; T38932; T38932.
DR GeneDB; Spombe; SPAC57A10.05C; --
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001880; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Ub1 conjugation; Repeat; WD Repeat.
FT DOMAIN 107 153
FT REPEAT 271 299 WD 1.
FT REPEAT 311 339 WD 2.
FT REPEAT 350 379 WD 3.
FT REPEAT 390 420 WD 4.
FT REPEAT 432 460 WD 5.
FT REPEAT 472 500 WD 6.
FT REPEAT 510 538 WD 7.
SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match
Best Local Similarity 19.5%; Score 590.5; DB 1; Length 605;
Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

```

QY 80 SSMIVPQKRLASVYKEKELCVKVFQOMSE---SDQVEFVEHLISQCHYQHGHINSYL 136
 Db 46 SSM-----HNEGLSGSEKQRVEAVMAAFSEASCSERKALQGLINNCSSLLSFASSTL 101
 QY 137 KPMQORDFTALPARGLDHIAENILSYLDKSLCAELVCKEYRVVTSQDGLMWKLIERM 196
 Db 102 DSLVRLDFLSLPV-----SISFRILSFUDARSLCQAQVSKHKELADDDVIWHRMCEQH 157
 QY 197 VRTDLSWRGLAERRGWGYLFKN----- 219
 Db 158 INRK-----CEKCGWGLPFLERNTLYAAKASIQRYERLTKRGVDQAHESPVKAKLD 211
 QY 220 -----KPPDGNAPPNSP-----YRALYKLIQDIETIESNWRGRHSLQ 258
 Db 212 DYTSSNEETISVKKPPSPNSKFLPKFTPKVEYVAERCR-----VECNWRHGR----- 263
 QY 259 RIHCR-----SETSGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECKRILTGHGTGVLCLQ 314
 Db 264 ---CROWLSGSHSDGVMCLQVRNLSAGSYDATIRLNLATFQOQVALLEGHSSGVTCCLQ 320
 QY 315 YDERVITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNNGMVTCSKDRSIADVDM 374
 Db 321 FQCKLISGSMKTRIMWYRSECSILHGTDSVLCCLTFDSTLVLGSGADCTVKLWHF 380
 QY 375 ASPTDITLRRVLVGHRAAVNVVDF--DDKYIVSASGDRITIKWNTSTCEPVRTLNGHKG 432
 Db 381 SGGKRITLR---GHTGPNVSVRIIRDRGLVLSGSDSTIKWSLETTNLTCLTFSAHIGP 436
 QY 433 IACLQVRDLVSGSSDNTIRLWDECGACLRVLGHEHLCVRCIRPDKRIVSGAYDGI 492
 Db 437 VQSLALADSLRFSCLSDGTIKQWIDIEKKCVHTLFGHIEGVMEIAADHLRLISGAHDGV 496
 QY 493 KWDVLAALDPRAPAGTCLRLTVEHSGRVFLQDFEQIVSSSHDDTILWDFLNDP 550
 Db 497 KWEEACE-----CVHTLKHSEPTSVVALGDCVSGSDGKIYLWLFNNAP 543

RESULT 7

SCOB_EMENI STANDARD; PRT; 678 AA.
 AC Q00659;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Sulfur metabolite repression control protein.
 GN SCOB OR MAPB1.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Natorff R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
 CC REPRESSION.
 CC -!- SIMILARITY: Contains 8 WD repeats.
 CC -!- SIMILARITY: Contains 1 F-box domain.
 CC -!- SIMILARITY: BELONGS TO THE MET30/SCOB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U21220; AAC15905.1; -;
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.

DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR ProDom; PD000018; WD40; 3.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS00682; WD_REPEATS_2; 7.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Transcription regulation; Repeat; WD repeat.
 FT DOMAIN 178 224 F-BOX.
 FT REPEAT 347 375 WD 1.
 FT REPEAT 387 415 WD 2.
 FT REPEAT 427 455 WD 3.
 FT REPEAT 466 496 WD 4.
 FT REPEAT 508 543 WD 5.
 FT REPEAT 553 595 WD 6.
 FT REPEAT 607 635 WD 7.
 FT REPEAT 647 675 WD 8.
 SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;
 Query Match 19.0%; Score 575; DB 1; Length 678;
 Best Local Similarity 26.9%; Pred. No. 7.4e-36;
 Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;
 QY 117 VEHLSIQMCHYGHINSYLPMLQRFITLPAAGLDHIAENILSYLDAKSLCAELVLC 176
 Db 153 LOGIWAQCCFPOLSVISATVRLIRIDITLPA---PEIAFKILCYLDTTSLCKASQVS 208
 QY 177 KEWYRVTSQDGLMWKLIERMVVRTDLSWRGLAERRGWGYLFKNK----- 220
 Db 209 RGWRALADDDVVMHRCQHIHRK-----CKKCGWGLPFLDRKRLRESKREIELRATTW 262
 QY 221 -----PPDQNA--PP-----NSFYRALY---P 237
 Db 263 DKGVVPASPSASBPSPGSKRKJEDDEVAVVXKHSLSGSDAGVDKSDPFKTRYPWK 322
 QY 238 KIIQDIETIESNWRGRHSLQRIHCRSETSK---GVYCLQYDDQKIVSGLRDNTIKWD 293
 Db 323 EYVKORFKVGTWNYGR-----CSIKTFKHTNGVMCLQFEDNLTATGSDYDTTIKW 375
 QY 294 KNTLECKRILTGHGTGVLCLQYDERVITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHL 353
 Db 376 TITGEELRTLGHESGIRCLQFDQDQTKLISGMDRTIKVMWRTGECISTYTGHRGGVIGL 435
 QY 354 RFNNGMVTCSKDRSIADVDMASPTDITLRRVLVGHRAAVNV--VDFDKYIVSASGDR 411
 Db 436 HFDAIILASGSDVKVKNWFEDKSTFSLR---GHTDWNANRVDTSSRTVFSASDDCT 491
 QY 412 IKWNTSTCEPVRTLNGH-----KRGIA----- 435
 Db 492 VRLWDLTKTLCIRTPHGHVGVQVVPVLPREFEFEEHDAECEDDLSTSGDANPPSIQA 551
 QY 436 -----LQYRD-----RLVVGSSDNTIRLWDIECGACLRVLEHLEL 472
 Db 552 SMGLPNAAYSQSSAFGTFPNGRAAPRYVWTSALDSTIRLWTTTGRCLRTFEGHLEG 611
 QY 473 VRCIRPDKRIVSGAYDGIKIVWDLVAALDPRAPAGTCLRLTVEHSGRVFLQDFEQI 532
 Db 612 VWALGADTLRIVSGAEDRMIKWD-----PRTGKERTFTGSGPVTICIGLGSRF 662
 QY 533 VSSSHDDTILWDF 546
 Db 663 ATGSEDCVEVRMYSF 676
 RESULT 8
 MT30_YEAST STANDARD; PRT; 640 AA.
 ID MT30_YEAST
 AC P39014;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)
 DT MET30 protein.
 EN MET30 OR Y11046W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=36089360; PubMed=8524217;
 RA Thomas D., Kuras L., Bathey R., Cherest H., Blaiseau P.L.,
 RA Surdin-Kerjan Y.
 RT "Met30p, a yeast transcriptional inhibitor that responds to S-
 RT adenosylmethionine, is an essential protein with WD40 repeats."
 RL Mol. Cell. Biol. 15:6526-6534(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169870;
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
 RA Meule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrall B.G.;
 RA "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX."
 RL Nature 387:84-87(1997).
 CC -|- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
 CC GENES EXPRESSION.
 CC -|- SUBUNIT: SEEMS TO INTERACT WITH MET4.
 CC -|- SIMILARITY: Contains 1 F-box domain.
 CC -|- SIMILARITY: Contains 8 WD repeats.
 CC -|- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC ENBL; Z46861; CAA86905.1; -;
 CC EMBL; L26505; AAA9671.1; -;
 CC PIR; S49932; S49932.
 CC SGD; S0001308; MET30.
 CC InterPro; IPR001810; F-box.
 CC InterPro; IPR001680; WD40.
 CC Pfam; PF00646; F-box; 1.
 CC PRINTS; PR00320; GPROTEINBRPT.
 CC ProDom; PD000018; WD40; 1.
 CC SMART; SM00256; FBOX; 1.
 CC SMART; SM00320; WD40; 6.
 CC PROSITE; PS50181; FBOX; 1.
 CC PROSITE; PS00678; WD_REPEATS_1; 4.
 CC PROSITE; PS00682; WD_REPEATS_2; 6.
 CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
 CC Transcription regulation; Methionine biosynthesis;
 CC Cysteine biosynthesis; Repeat; WD repeat.
 KW DOMAIN 181 227 F-BOX.
 FT REPEAT 300 328 WD 1.
 FT REPEAT 340 368 WD 2.
 FT REPEAT 380 408 WD 3.
 FT REPEAT 419 449 WD 4.
 FT REPEAT 461 499 WD 5.
 FT REPEAT 509 538 WD 6.
 FT REPEAT 550 578 WD 7.
 FT REPEAT 607 635 WD 8.
 FT REPEAT 61 61 M -> I (IN REF. 1).
 FT CONFLICT 61 61
 SQ SEQUENCE 640 AA; 72835 MW; 5135D4BCA2E1B97 CRC64;
 Query Match 18.0%; Score 545; DB 1; Length 640;
 Best Local Similarity 28.9%; Pred. No. 1.3e-13;
 Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;
 QY 34 IPEKNSLRQTYNSCARCLN-QETVCLASTAMKTEVCVAK-TKLANGTSSMIVPKORKLS 91
 DB 86 LPEYNFTKFCYRHNPDIOFPTHTACYPKQDLKRTQINAKIAKPLQEQSDIHHIISKYS 145
 QY 92 ASYEKEKELCVKYPEQWSESDQVEFVEHLISQCHYGHINSYLPKMLQORDEITAPAR 151
 DB 146 NSNDKIRKL-----LDGILSTSCFPQLSYISLVTHMKIDFISILP- 188
 QY 152 GLDIAENILSYLDAKSLCAELVCKEYRVTSGLMKLIERMVRTDSLRLGLAERR- 210
 DB 189 --QELSLKILSYLDCQLCNATRCRKWKQLADDDRVVYHMCBOHI-----DRKC 236
 QY 211 ---GWGQVLFKNK-----PPDGNAPNSFYRALYPKIOTDIETIESNNRCGHSILQ 258
 DB 237 PNCWGGLPLHMKRARIQNSTGSSNADIQTTPRPWKVIYRERPKVESNWRKG----- 291
 QY 259 RIHCRSETSK----GVYCLQYDQKIVSGLRDNTIKIDKNTLECKRIILTGTGSLCLQ 314
 DB 292 --HCRIQEFKGMGVLTLOFNVELLFTGSDYSTIGWDLFTGKLRRLSGHSDGVKTLY 349
 QY 315 YDERVLIITGSSDSIVRVWDVNTGEMNLTLHHCBAVLHLPFNNGMMVTCKSKRSTAVDM 374
 DB 350 FDDRKLITGSLDKTIRVNNYITGECISTYRGHSDSVLSVDYQKIVSGSADTKVKVMEV 409
 QY 375 ASPDTIDILRRVLVGHRAAVNVVDFOKYI--VSGDRTIKVWNTSTCFVRTLNGH--- 429
 DB 410 ESRICYTLR---GTEWVNCVKLHPKSPFCSCDDTIRWMDITNSCLKVFRGHVQ 465
 QY 430 KEGTACLOYRP--RLVSSGSS----- 448
 DB 456 VKIIPITKIDVENLATDNTSDGSSPODDTMDGADSDTPSNEQETVLDENIPYPTHL 525
 QY 449 -----DNTIRLWDECCACLRVLEGHEELVRCIRFONKRVISGAYGKIKWDLVAALDP 503
 DB 526 LSCGLDNTIKLWDVTKCKIRTPQGHVGVWDIADNFRIISGSHDGSIKVWDLOSG--- 582
 QY 504 RAPAGTCLRLTVHSGHVRFLQDFEQIVSSSHDDTI 541
 DB 583 -----KCMHTF---NGR--RLQRETOHTQTQSLGDKV 609
 RESULT 9
 SC02_NEUCR
 ID SC02_NEUCR STANDARD; PRT: 650 AA.
 AC Q01277;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Sulfur controller-2 (SC02).
 GN SC02-2.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 CX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A / FGSC 987;
 RX MEDLINE=95241499; PubMed=7724564;
 RA Kumar A., Paletta J.V.;
 RT "The sulfur controller-2 negative regulatory gene of *Neurospora*
 RT crassa encodes a protein with beta-transducin repeats."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
 CC -|- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
 CC -|- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
 CC -|- SIMILARITY: Contains 8 WD repeats.
 CC -|- SIMILARITY: Contains 1 F-box domain.
 CC -|- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.

116 FVEHLISOMCHYQHGHINSYKPMQLORDPITALPARGLDHIAENILSYLDAKSLCAAEV 175
 87 AFTELLQESNWTNIRQLRAIEPHFORFLSLCPV---ELGKMLHNLTCYDLKVAQV 142
 176 CKRYRVTSOGMLWKL-TERMV-----RDSLRWGLAERGGVLFKNKPPDGNAP 227
 143 SKWKLLISEDKIWKSLGVBEFFKHHPDPTDRVTGAWOGTAAG-----VTIPDHQP 195
 228 PN-SFYRALYPKIIQDI-----ETIESNRCGRHSRLORICRSETSGVYCL 273
 196 CDLNVHFLKQFGDIFERAADSKSYLRADKIEKNWNPIMPISAV-LRGHEDHVTICM 254
 274 QYDOKIVSGLRDNITKWDKNTLECKRILTGHTGSLVLOYDE--RVIIITSGSDSTVRV 331
 255 QIHDDVLVTGSDNTLKWICIDKEVYMLVHGHTGVTWSQISQCGRIYVSGSDTRTKV 314
 332 WDVNTGMLNLTLIHCEAVLHLEFNNGMVTCRSKORSIAVMDMAFPTDITLRRVLVGHRA 391
 315 WSTVDGSLHTLQHTSTVRCMAMAGSILVTGSRDITLRVMDVESGRHLA---TLHGHA 371
 392 AVNVVDFDDKYIVSASGDRTIKWNNTSTCFVRTLNHKGKGIACLOYDR--LVVSGSSD 449
 372 AVRCVQDGTIVVSGGVDFVTKVNAHTGRCIRTLTGHNRRVYSLFESERSIVCSGSLD 431
 450 NTRLWDI---ECGACLRVLEGEHELVRICIRFDNKRIVSGAYDGIKIVMDLVAALDPRAP 506
 432 TSIRVMDFTPEGQECVALLQHTSLTSGMLRGNILVSCNADSHRVWDI-----H 483
 507 AGTLCLATLVEHSGRVRLQD-FEFOIVSSSHDDTLIWD 545
 484 EGT-CVHMLSGHRSATLSLQWFRGNRVATSSDDGIVKLWD 522

RESULT 11
 CDC4 CANAL STANDARD; PRT; 684 AA.
 ID CDC4 CANAL
 AC P53699
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cell division control protein 4.
 EN CDC4.
 OS Candida albicans (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 CC NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SGY126;
 RA Shieh J.C., White A.M., Rosamond J.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
 CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
 CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
 CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
 CC VARIOUS ASPECTS OF SPOULATION. REQUIRED FOR HTAL-HTBI LOCUS
 CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).
 CC -!- SIMILARITY: Contains 1 F-box domain.
 CC -!- SIMILARITY: Contains 7 WD repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X96763; CAA65538.1; -
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00q46; F-box; 1.
 DR Pfam; PF00400; WD40; 7.

PRINTS; PRO0320; GPROTEINBRPT.
 DR SMART; SMO0256; FBOX; 1.
 DR SMART; SMO0320; WD40; 7.
 DR PROSITE; PS0181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS0082; WD_REPEATS_2; 4.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
 FT DOMAIN 212 258 F-BOX.
 FT REPEAT 322 351 WD 1.
 FT REPEAT 363 391 WD 2.
 FT REPEAT 403 431 WD 3.
 FT REPEAT 442 468 WD 4 (POTENTIAL).
 FT REPEAT 478 506 WD 5.
 FT REPEAT 519 549 WD 6.
 FT REPEAT 561 589 WD 7.
 SQ SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;
 Query Match 15.7%; Score 477.5; DB 1; Length 684;
 Best Local Similarity 25.4%; Pred. No. 1.8e-28;
 Matches 136; Conservative 95; Mismatches 184; Indels 121; Gaps 16;
 QY 109 SESQVFEVHLISOMCHYQHGHINSYKPMQLORDPITALPARGLDHIAENILSYLDAKS 168
 DB 179 SEANQNLVFLKLOKTRPTLTSTFNLLINSLKRDILSNVFP---EVTMKILSYLDYKT 234
 QY 169 LCAAEVLCKEYRVTSOGMLWKLIER--MVRTDSLWRGLAERGGVLFKNKPPD--- 223
 DB 235 LLSVAQVCKKWFIDINNPDPDWIKLXKDKLITDDAVIK-----VELQYPDQLL 282
 QY 224 ---GNAPNSFYRALYPKIIQDIETIESNRCGRHSRLORICRSETSGVYCLQYDDOKI 280
 DB 283 REWSTLPEINSAQVLYKK---RKIIIVNRMDPKFKPHRISVSGHGNKVVTCLQHDDEKV 338
 QY 281 VSGLRDNITIKWDKNTLECKRILTGHTGSLVLOYDERVVIITGSSDSTVRVMDVNTGML 340
 DB 339 VTGVDDRCISIVSTQTQOLMKVLEGGVWALKYTGNTLVGTSTDRTVRVNMTGTCT 398
 QY 341 NTLIHH---CEAVLHL---RPNNG-----MMVTCSDRSIAVDM----- 374
 DB 359 HIFRGHTSTIRCLDIHHPAIVGKNQDGEIVPEYFELLITGSRDHNHVMKLVFVVDSDQ 458
 QY 375 -----ASPTDITLR-----RVLV 387
 DB 459 VIETPEGEFDPNFIYLVLSGHTQSVRSISGVGNIIFSGSYDSTVRVWDLDDGHCHTVLQ 518
 QY 388 GH--RAAVNVVDFDDKYIVSASGDRTIKWNNTSTCFVRTLNHKGKGIACLOYRDLVVS 445
 DB 519 GHLDRVYSTAIDFHSKTCFSGMSDININWNFTGELKKVLVGHASLVGLGLDLVDVVLVS 578
 QY 446 GSSDNTIRLNDIECGACLRVLEGEHELVRICIRFDNKRIVSGAYDGIKIVMDLVAALDPR 505
 DB 579 AAADATLIRWDAKTGELRSKLKGHGAATCFEHDGLRVVSGS-EKMLKLVN----- 629
 QY 506 PAGTILCLATLVEHSGRVRLQD-----EFOIVSSSHDDTLI-IMDFLNDPAAQ 553
 DB 630 EKGTFARDLLSDVTGGIWDQVRFDYKRCVAAVQRIINEDEGETFIELDP-SQPLIQ 684

RESULT 12
 HET1 PODAN STANDARD; PRT; 1356 AA.
 ID HET1 PODAN
 AC Q00808;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vegetatable incompatibility protein HET-E-1.
 GN HET-E1.
 OS Podospora anserina.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.
 CC NCBI_TaxID=5145;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=96009891; PubMed=7557402;
 RA Saue S., Turcq B., Begueret J.;
 RT "A gene responsible for vegetative incompatibility in the fungus
 RT Podospira anserina encodes a protein with a GTP-binding motif and G
 RL beta homologous domain.";
 RL Gene 162:135-139(1995).
 CC -!- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
 CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
 CC HET-C.
 CC -!- SIMILARITY: Contains 10 WD repeats.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L28125; AAA85775.1; -;
 DR PIR; T18521; T18521.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 10.
 DR PRINTS; PR00320; GPROTEINRPT.
 DR ProDom; PD000018; WD40; 10.
 DR SMART; SM00320; WD40; 10.
 DR PROSITE; PS00837; NACHT; 1.
 DR PROSITE; PS00678; WD REPEATS 1; 10.
 DR PROSITE; PS00082; WD REPEATS 2; 10.
 DR PROSITE; PS02094; WD REPEATS REGION; 1.
 KW GTP-binding; Repeat; WD repeat.
 FT DOMAIN 294 629 NACHT.
 FT NP_BIND 300 307 GTP (POTENTIAL).
 FT REPEAT 839 869 WD 1.
 FT REPEAT 881 911 WD 2.
 FT REPEAT 923 953 WD 3.
 FT REPEAT 965 995 WD 4.
 FT REPEAT 1007 1037 WD 5.
 FT REPEAT 1049 1079 WD 6.
 FT REPEAT 1091 1121 WD 7.
 FT REPEAT 1133 1163 WD 8.
 FT REPEAT 1175 1205 WD 9.
 FT REPEAT 1217 1247 WD 10.
 SQ SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;
 Query Match 15.0%; Score 455.5; DB 1; Length 1356;
 Best Local Similarity 31.7%; Pred. No. 26-26;
 Matches 121; Conservative 64; Mismatches 131; Indels 66; Gaps 15;
 QY 226 APPNSFYRLY-----PKIIQDIETIESNRCGRHSL-----QRIHCRS- 264
 DB 802 APTDSMTIKKXEPGWNISTISVVEAEVNACTQTLEHGSSVLSVAFSADGQVAFSGSD 861
 QY 265 -----ETSG-----VYCLOY--DDQKIVSLRNTIKWTKLTKRIL 303
 DB 862 DKTIKWDTAGTQTLEHGSSVWVAFSPDRERVASGSDDKTIKIWDAAAGTCTQTL 921
 QY 304 TGTGTVSLCLOY--DERVIITGSDSTVRVVDVNTGEMTLTIHCEAVLHFRN-NGNM 360
 DB 922 EGHGRGVSVAFSPDQGVASGSDDKTIKIWDAAAGTCTQTLEHGSSVLSVAFSPDQGR 981
 QY 361 VTC-SKRSIAVMDASPTDILRLVLGHRAAVVVDV--DKYIVVSGRDTIKWNT 417
 DB 982 VASGSDTKIKWDTAG--TCTQTLEHGSSVWVAFSPDQGVASGSDDKTIKIWD 1038
 QY 418 STCEPVRTLNHGKRGIACTQYR--DRLVWGSNDNTIRLWDIECGACLRVLGHEELVRC 475
 DB 1039 ASGCTQTLEHGSSVWVAFSPDQGVASGSDDKTIKIWDVAGTCTQTLEHGSSVWS 1098
 QY 476 IRF--DNKRIVSGYDGIKIVMDLVAALDPAPAGTLCRLTVLHSGRVRLOF--DEFO 531

Db 1099 VAFSPDQGVASGSDGTGIKIWD-----AASGT-CTQTLEHGSSVWVAFSPDQGR 1149
 QY 532 IVSSSHDITLWDFLNDPAQAQ 553
 Db 1150 VASGSDGTGIKIWDAAAGTCTQ 1171
 RESULT 13
 POP1 SCHPO
 ID POP1 SCHPO STANDARD; PRT; 775 AA.
 AC P37060; Q9P7P3;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE WD-repeat protein popl.
 GN POP1 OR SPBC1718.01 OR SPBC2G2.18.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97347242; PubMed=9203581;
 RA Kominami K., Toda T.;
 RT "Fission yeast WD-repeat protein popl regulates genome ploidy through
 RT ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rml1
 RT and the S-phase initiator Cdc18.";
 RL Genes Dev. 11:1548-1560(1997).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolckaert G., Aert R., Robben J., Grymonprez B.,
 RA Woltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 RN [3]
 RC SUBUNIT.
 RC STRAIN=972;
 RX MEDLINE=99144318; PubMed=9990507;
 RA Kominami K.-I., Ochofrena I., Toda T.;
 RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
 RT complexes together with cullin-1 in fission yeast SCF (Skip-
 RT cullin-1-F-box) ubiquitin ligase.";
 RL Genes Cells 3:721-735(1998).
 RN [4]
 RP FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 RX MEDLINE=22325332; PubMed=12167173;

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X05625; CAA29113.1; -
 CC EMBL; D50617; BAA09229.1; -
 CC EMBL; Z46255; CAA86341.1; -
 CC PIR; S56245; S56245.
 CC PDB; INEX; 18-FEB-03.
 CC SGD; S000185; CDC4.
 CC GO; GO:000082; P-G1/S transition of mitotic cell cycle; IMP.
 CC GO; GO:000086; P-G2/M transition of mitotic cell cycle; IGI.
 CC InterPro; IPR001810; F-box.
 CC Pfam; PF00646; F-box; 1.
 CC Pfam; PF0400; WD40; 6.
 CC PRINTS; PR00320; GPROTEINRPT.
 CC ProDom; PD000018; WD40; 1.
 CC SMART; SM00256; FBOX; 1.
 CC SMART; SM00320; WD40; 7.
 CC PROSITE; PS0181; FBOX; 1.
 CC PROSITE; PS0678; WD REPEATS; 1; 4.
 CC PROSITE; PS0082; WD REPEATS; 2; 5.
 CC PROSITE; PS0294; WD REPEATS REGION; 1.
 CC Cell division; Mitosis; Sporulation; Repeat; WD repeat; 3D-structure.
 CC DOMAIN 272 319 F-BOX.
 CC REPEAT 380 408 WD 1.
 CC REPEAT 420 449 WD 2.
 CC REPEAT 461 493 WD 3.
 CC REPEAT 528 556 WD 4.
 CC REPEAT 568 598 WD 5.
 CC REPEAT 630 658 WD 6.
 CC REPEAT 669 698 WD 7.
 CC CONFLICT 460 460 K -> E (IN REF. 1).
 CC SEQUENCE 779 AA; 86089 MW; 0348F2F8FA78F3BC CRC64;
 Query Match 13.23; Score 399; DB 1; Length 779;
 Best Local Similarity 24.33; Pred. No. 1.8e-22;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;
 QY 56 TVCLASTAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEK-----ELCVKYPQW 108
 DB 194 TPLAKTKTKINN-----NNNIADLESKDSIIISPEYLSDEISAINNLPHAYPK-- 244
 QY 109 SESQVEFEVHLISQCHYQGHGINSYKPLQDEITAPLARGLDHAEILSLYDRAKS 168
 DB 245 -----NLLFLVANMORSELSGLTIKLNKRDLTISLFP-----EISLKFNYLQPED 294
 QY 169 LCAAEALYCKEYR-VTSDGMLWKKLI--ERMVRTDSLWGLAERRGQGYLFKNKPPDGN 225
 DB 295 IINSLGVSNQNKTKIRKSTSLWKKLIISNFV-----SPKGF 331
 QY 226 APPNSFYBALYFKLIQD-----IET--TESNWRGRLSHLRIHCRSETSGVYVCLOYD 276
 DB 332 NSLNKLSQKPKLSQQRRLSLFENFLKWNYPKFPVORTLIRGHMTSVITCLQFE 391
 QY 277 DQKIVSGLRDNTIKWDRNTLECKRILTGTGSLVCLQYDE-RVITGSSDSSTVRVNDVN 335
 DB 392 DNYVITGADRRVIRVYDSINKKFLQLSGHDGGVWALKYAHGGLIVSGSDTRTVRVNDIK 451
 QY 336 TGEMLNTLIHCEAVLHLRFNNGMWTCSKDRSIADVNDMASTDITLRVLVGHAAVNV 395
 DB 452 KG-----CCT-----HVFKGHNSTVRC 468
 QY 396 VDFDD-----KYIVASGDRTIKVNWT-----STCE-----FVRLTNG 428
 DB 469 LDIVEYKNIKIVTGSRDNTLHVWKLPKSESSVPDHGEEDYPLVFHTPEENPYFVGVLRG 528
 QY 429 HKRGIAQLQYEDRVLVSGSSDNTIRWDIECGACIYVLEGHEELVRCIRFDN--KGIYSG 486
 DB 529 HMASVTVSGHNVISGSDYNTLIVWDAVMQMKVLISGHTDRIVSTIYDHERKRCISA 588

QY 487 AYDGKIKVWDL-----VAALDPRAGTU--CLRTLVEHSGRVFRLQDFEQIVSS 535
 DB 589 SMDTIRLWDLNENWNGESYATNSAPKILGMYTLOGHIALVGLRLSLRSLFVSA 648
 QY 536 SHDITLWDLNPDPAQAQPPRSPRTYTY 566
 DB 649 AADGSIKWD-AND-----YSRKFSY 668
 RESULT 15
 POP2 SCHPO
 ID POP2 SCHPO STANDARD; PRT; 703 AA.
 AC 014170;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE WD-repeat protein pop2 (Proteolysis factor sud1).
 GN POP2 OR SUD1 OR SPAC47.03.
 OS Schizosaccharomyces pombe (Pission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=99227353; PubMed=10209119;
 RA Wolf D.A., McKeon F., Jackson P.K.;
 RT "F-box/WD-repeat proteins pop1p and Sud1p/Pop2p form complexes that
 RL bind and direct the proteolysis of cdc18p.";
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98318628; PubMed=9653157;
 RA Jallepalli P.V., Tien D., Kelly T.J.;
 RT "sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1
 RT proteins for degradation and stops unwanted diploidization in fission
 RT yeast.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=2184401; PubMed=11659360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sprouks J., Feat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth I., Churcher C.N.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren I., Whitehead S.,
 RA Woodward J., Voicakart G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Gadiu S., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.R., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 RN [4]
 RP FUNCTION, AND SUBUNIT.
 RC STRAIN=972;
 RX MEDLINE=99144318; PubMed=9990507;

RA Kominami K.-I., Ochotorena I., Toda T.;
 RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
 RT complexes together with cullin-1 in fission yeast SCF (Skip-
 RT cullin-1-F-box) ubiquitin ligase.";
 RL Genes Cells 3:721-735(1998).
 RN [5]
 RP FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 RX MEDLINE=22325332; PubMed=12167173;
 RA Seibert V., Prohl C., Scholtz I., Rhee E., Lopez R., Abderazzaq K.,
 RA Zhou C., Wolf D.A.;
 RT "Combinatorial diversity of fission yeast SCF ubiquitin ligases by
 RT homo- and heterooligomeric assemblies of the F-box proteins Pop1 and
 RT Pop2p.";
 RL BMC Biochem. 3:22-22(2002).
 CC -!- FUNCTION: Involved in maintenance of ploidy through proteasome
 CC dependent degradation of CDK inhibitor rum1 and S-phase initiator
 CC cdc18. Functions as a recognition factor for rum1 and cdc18, which
 CC are subsequently ubiquitinated and targeted to the 26S proteasome
 CC for degradation.
 CC -!- SUBUNIT: Homodimer and heterodimer with pop1. Binds to pcu1, pip1
 CC and cdc18.
 CC -!- SUBCELLULAR LOCATION: Nuclear, and cytoplasmic.
 CC -!- SIMILARITY: Contains 1 F-box domain.
 CC -!- SIMILARITY: Contains 6 WD repeats.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; AF038867; AAB95480.1; -
 CC DR ENBL; AF064515; AAC39496.1; -
 CC DR ENBL; Z98602; CAB11275.1; -
 CC DR PIR; T43557; T43557.
 CC DR GeneDB Spombe; SPAC4D7.03; -
 CC DR InterPro; IPR001810; F-box.
 CC DR InterPro; IPR001680; WD40.
 CC DR Pfam; PF00646; F-box; 1.
 CC DR Pfam; PF00400; WD40; 6.
 CC DR PRINTS; PR00320; GPROTEINERPT.
 CC DR PRODOM; PD000018; WD40; 1.
 CC DR SMART; SM00256; FBOX; 1.
 CC DR SMART; SM00320; WD40; 7.
 CC DR PROSITE; PS50181; FBOX; 1.
 CC DR PROSITE; PS00678; WD_REPEATS_1; 3.
 CC DR PROSITE; PS50082; WD_REPEATS_2; 6.
 CC DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 CC KW Nuclear protein; Repeat; WD repeat.
 CC FT DOMAIN 236 283 F-BOX.
 CC FT REPEAT 389 417 WD 1.
 CC FT REPEAT 429 473 WD 2.
 CC FT REPEAT 505 533 WD 3.
 CC FT REPEAT 545 575 WD 4.
 CC FT REPEAT 587 615 WD 5.
 CC FT REPEAT 625 654 WD 6.
 CC SQ SEQUENCE 703 AA; 79605 MW; 0336A0568C152396 CRC64;
 Query Match 13.1%; Score 396.5; DB 1; Length 703;
 Best Local Similarity 23.3%; Pred. NO. 2.4e-22;
 Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;
 QY 14 KFMNRRERD-CNNGEPPPKI-----IPKNSLRQYNSCARLCL----NQETVCL 59
 DB 96 KFNESIVSVKLDHSPDPSVPGDFSIILVFORNPL-YSHSLLPKIISIDRNRIKLD 154
 QY 60 ASTAMKTENCVAKTKLANGTSSMIVPKORKLSAYEK-EKELCVKYFFQWSESQVEFVE 118
 DB 155 NSISNSDNFPSPKPV--DTSNTVSPGSPISLEDLEDLNLQSIQVTFEDLPFGIQ-SYAP 211
 QY 119 HLIQCMCHYQGHG-INSYLKPLQDFITALPARGLDHAENILSYLDAKSLCAELVCK 177

DB 212 FQLRSCNRSOMLLNECEPLKKDILNSLPF-----SIVOSILLNLDIHSLFSLCVSP 267
 QY 178 EWRV-----TSDGMLWKXLIERMVRTDSLWRGLABRRGQYLFKNKPPDGNAPP---N 229
 DB 268 TWRILDVHTS---YWKHMF-----SLFGQIENENDW-----KYANFNLRPPFLHND 312
 QY 230 SFYEALVPKIIQDIETIESNWRGSRHSIOR-----
 DB 313 QISDDYPEIFK-----RHFLNRKRWLPSPPPSHLSFPPIHVPNFMITSLLH 360
 QY 260 -----IHKSETSK-GVYCLOYDQKIVSGSLRDNTIKIMDKNT 296
 DB 361 KDIRITTSGGTTIHNATIGVLEARLEHGEKGVAVKIHENTLVSGSIDKTVRVNIEK 420
 QY 297 LECKRIITHTGTVLCLOY-----DERVITGSSDSSTVRVWDVNTGEML 340
 DB 421 AKCTHIFEGHISIRCEILVPSRLIRHGVEIVEPDQPIYVSGSRDHTLRVW----- 472
 QY 341 NTLIHCEAVLHFRFNNGMVTCSDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDD 400
 DB 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNPYFV---HTLIGHTSDVRTISYG 517
 QY 401 KYIVSASGDRTIKVMNTSTCEFVRLNGHKRGCIACLOYRD--RLVVGSSDNTLWLDIE 458
 DB 518 DILVSGSYDSIRKINRVSTGECYLHURGHSIRIYSLVEPERNICISGMDKSIKRVWDL 577
 QY 459 CGACLRVLEGGHEELVRCIRFDNKNRIVSGAYDGKIKVWDLVAALDPAPAGTLCRLTVEH 518
 DB 578 TGTCKYVLEGHDAFVTLNLFVQNRLIISGSADSTIRWD---LNTGKP-----LMVLPSN 628
 QY 519 SGRVRLQDFEFQIVSSSHDDITLIWD 545
 DB 629 SGYISSFVSDEHKII-SGNDGSKLMD 654

Search completed: October 22, 2003, 10:23:22
 Job time : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

1M protein - protein search, using sw model

Run on: October 22, 2003, 10:06:36 ; Search time 84 seconds

(without alignments)
1075.184 Million cell updates/sec

File: US-09-601-168B-2

Perfect score: 3034
Sequence: 1 MDPABAVLQKALFKMNSSE.....PAAQAEPPSPSTYTYISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3034	100.0	569	AA124054	A human beta-trans
2	3034	100.0	569	AA12813	Human beta-transdu
3	3034	100.0	569	AA196697	Human beta-TrCP.
4	3034	100.0	569	AA183041	F-box protein FBP-
5	3034	100.0	569	AA183250	F-box protein hbet
6	3034	100.0	569	AA183250	Human cell signal
7	3034	100.0	569	AA144249	Human F-box protein
8	3034	100.0	569	ABG69473	Human bait protein
9	3027	99.8	569	AA148298	Human Fli1 protein

10	3027	99.8	608	22	AA100960	Human bone marrow
11	3066	99.1	605	22	AA178582	Human protein SEQ
12	2997	98.8	569	21	AA12813	Mouse ubiquitin li
13	2997	98.8	569	21	AA183254	F-box protein FWDI
14	2992.5	98.6	632	22	AA178584	Human protein SEQ
15	2989.5	98.5	654	22	AA179566	Human protein SEQ
16	2989.5	98.5	654	22	AA179567	Human protein SEQ
17	2989.5	98.5	654	22	AA179568	Human protein SEQ
18	2962	97.6	590	22	AA100847	Human bone marrow
19	2882	95.0	579	22	AA178583	Human protein SEQ
20	2882.5	95.1	517	16	AA183582	WD-40 domain-contg
21	2884.5	78.6	542	21	AA196696	Human E3 ubiquitin
22	2884.5	78.6	542	22	AA179127	Human protein SEQ
23	2884.5	78.6	542	22	AA140208	Human polypeptide
24	2884.5	78.6	550	22	AA141994	Human polypeptide
25	2089.5	68.9	510	22	AA159857	Drosophila melanog
26	1163	38.3	219	23	AA198087	Human beta TrCP (b
27	989.5	32.6	265	22	AA186942	Human DNA repair a
28	935	30.8	448	22	AA182839	Beta-TrCP.N/SKP2.C
29	661	21.8	1326	22	AA1867237	Drosophila melanog
30	661	21.8	1326	22	AA1867238	Drosophila melanog
31	661	21.8	1326	22	AA170051	Drosophila melanog
32	644.5	21.2	626	20	AA122469	Human myc-N-sei-1
33	644.5	21.2	626	22	AA159201	Protein encoded by
34	640	21.1	540	20	AA122465	Human hippocampal
35	640	21.1	540	20	AA122468	Human mammary sel-
36	640	21.1	540	22	AA159197	Human hippocampal
37	640	21.1	540	22	AA159200	Human mammary sel-
38	640	21.1	545	20	AA122464	Human hippocampal
39	640	21.1	545	22	AA159196	Human hippocampal
40	640	21.1	553	22	AA122463	Human hippocampal
41	640	21.1	553	22	AA193475	Human protein sequ
42	640	21.1	553	22	AA159195	Human hippocampal
43	640	21.1	559	20	AA122467	Human mammary sel-
44	640	21.1	559	22	AA159199	Human mammary sel-
45	640	21.1	589	20	AA122466	Human mammary sel-

ALIGNMENTS

RESULT 1
AA124054
ID AA124054 standard; Protein; 569 AA.

XX AC AA124054;

XX 20-MAR-2003 (updated)

DT 30-SEP-1999 (first entry)

XX A human beta-transducin repeat containing protein.

XX Beta-transducin repeat containing protein; beta-TrCP; Skp1p;
XX proteasome degradation pathway; Vpu protein; beta-catenin;
XX human immune deficiency virus-1; HIV-1; cellular protein; IkappaB;
XX ubiquitination; phosphorylated protein; tumour; apoptosis; Alzheimer's;
XX antiviral; antitumour; cell cycle regulation; protein degradation;
XX and anti-inflammatory; osteo-articular inflammation; acute inflammation;
XX tumour necrosis factor.

XX Homo sapiens.

Key	Location/Qualifiers
Region	147..191
Region	/note= "F box sequence"
Region	259..292
Region	/note= "WD motif"
Region	304..332
Region	/note= "WD motif"
Region	343..372
Region	/note= "WD motif"
Region	387..415
Region	/note= "WD motif"

Region 427..455
/note= "WD motif"
Region 467..492
/note= "WD motif"
Region 516..544
/note= "WD motif"

409338969-A1.

05-AUG-1999.

29-JAN-1999; 99WO-PR00196.

29-DEC-1998; 98FR-0015545.

0-JAN-1998; 98FR-0001100.

INSP } INST PASTEUR.

INRM } INST NAT SANTE & RECH MEDICALE.

renzana Seidedos F, Benarous R, Concordet J, Durand H;

Rolli M, Margottin F;

PI; 1999-469329/39.

-PSDB; AAX86501.

ew human beta-transducin repeat containing protein and its
ragments useful as, or to screen for, antiviral, antitumour,
nti-inflammatory and anti-Alzheimer's agents

laim 1; Page 60-61; 71pp; French.

he present sequence represents a human beta-transducin repeat containing
rotein (beta-TrCP). The protein directs proteins to the proteosome
agradation pathways. The protein is able to interact with the Vpu
rotein of human immune deficiency virus-1 (HIV-1), cellular proteins
appab or beta-catenin (bc) and/or protein Skp1. The protein controls
liquinylation of phosphorylated proteins and thus their targeting to
ocosomes for degradation. Depending on whether the process is
inhibited or promoted, the result may be delayed breakdown of CD4 (in
ses of HIV-1 infection); increased activity of Ikb (and thus reduced
tivity of NFkappaB) and increased degradation of mutant bc in tumour
lls, or increased bc survival (and reduced apoptosis) in Alzheimer's
tients. The beta-TrCP protein, and its active peptide fragments, or its
cleic acid, are used to screen for anti HIV-1 agents (antivirals),
t tumour agents that disrupt cell cycle regulation or protein
gradation in human tumour cells, and anti-inflammatory agents that
rupt activation by NFkappaB. Fragments of the protein are also
etui for treating osteo-articular inflammation or acute inflammation
sociated with release of tumour necrosis factor.
dated on 20-MAR-2003 to correct PA field.)

quence 569 AA;

Match 100.0%; Score 3034; DB 20; Length 569;
Local Similarity 100.0%; Pred. No. 2.5e-286;
es 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPAEAVLQEKALPKFMSNSEDNCNGEPKRIPEKNSLRQTYNSCARLCLNQSTVCLA 60

1 MDPAEAVLQEKALPKFMSNSEDNCNGEPKRIPEKNSLRQTYNSCARLCLNQSTVCLA 60

61 STAMKTEVCVAKTKLANGTSMIVPKOKLSASVEKEKELCVKYPFQWSESQVFEVHL 120

61 STAMKTEVCVAKTKLANGTSMIVPKOKLSASVEKEKELCVKYPFQWSESQVFEVHL 120

121 ISQMCHYQGHINSYKPMPLQDFTIPALPARGLDHIAENTLSYLDKASLCAAELVCKEWY 180

121 ISQMCHYQGHINSYKPMPLQDFTIPALPARGLDHIAENTLSYLDKASLCAAELVCKEWY 180

181 RVTSDGLWKKLIERMVTDSLWGLAERRGNGGYLPKKNPPDGNAPPNSFYRALYPKII 240

181 RVTSDGLWKKLIERMVTDSLWGLAERRGNGGYLPKKNPPDGNAPPNSFYRALYPKII 240

QY 241 QDIETIESNWRGCRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
DB 241 QDIETIESNWRGCRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
QY 301 RILTGHGTSVLCLOYDERVITIGSSDSSTVVRVMDVNTGEMWNTLIHHCEAVLHLPFNNGMM 360
DB 301 RILTGHGTSVLCLOYDERVITIGSSDSSTVVRVMDVNTGEMWNTLIHHCEAVLHLPFNNGMM 360
QY 361 VTCSDKRSIAVMDVASPTDITLRRVLVGHRAAVNVVDFDKKYIVSASGDRTIKVMNTSTC 420
DB 361 VTCSDKRSIAVMDVASPTDITLRRVLVGHRAAVNVVDFDKKYIVSASGDRTIKVMNTSTC 420
QY 421 EFVETLNHKGRIACLOYRDLRVVSGSSDNTTLRLWDIECGACLRVLEGHEELVRCIRFDN 480
DB 421 EFVETLNHKGRIACLOYRDLRVVSGSSDNTTLRLWDIECGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCLRTLVEHSGRVRFLQDFEQIVSSSHDDT 540
DB 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCLRTLVEHSGRVRFLQDFEQIVSSSHDDT 540
QY 541 ILIWDFLNDPAAQAEPPSPSRITYTISR 569
DB 541 ILIWDFLNDPAAQAEPPSPSRITYTISR 569

RESULT 2

AAB12813

ID AAB12813 standard; protein; 569 AA.

AC AAB12813;

XX 27-NOV-2000 (first entry)

DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.

XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;

XX beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;

XX Gene therapy; colon cancer; beta-transducin repeat containing protein;

XX beta-TrCP.

OS Homo sapiens.

XX JP2000166542-A.

XX 20-JUN-2000.

XX 02-DEC-1998; 98JP-0343437.

XX 02-DEC-1998; 98JP-0343437.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2000-485550/43.

XX N-PSDB; AAA73132.

XX F-box protein of ubiquitin ligase SCF complex which promotes the

XX ubiquitination of IkappaB or beta-catenin

XX Claim 3; Page 10-12; 19pp; Japanese.

XX The present invention describes an F-box motif protein of ubiquitin
CC ligase SCF complex which promotes the ubiquitination of IkappaB or
CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
CC complex (SCF complex) of F-box protein containing F-box motif and WD40
CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)
CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin
CC repeat containing protein (beta-TrCP)). The F-box protein can be used for
CC the gene therapy of colon cancer by being recombinated to a virus vector.

XX Sequence 569 AA;

Query Match 100.0%; Score 3034; DB 21; Length 569;

Local Similarity 100.0%; Pred. No. 2.5e-286; Indels 0; Gaps 0;
Res 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPAAVLQEKALFNMSSERDCNNGEPKRIIPEKNSLRQTYNSCARCLNQETVCLA 60
1 MDPAAVLQEKALFNMSSERDCNNGEPKRIIPEKNSLRQTYNSCARCLNQETVCLA 60
61 STAMKTENCVAKTTLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
61 STAMKTENCVAKTTLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
121 ISQCHYQHGHSINLYKPMQLQROFITALPARGLDHIAENILSYLDAKSLCAELVCCKEWY 180
121 ISQCHYQHGHSINLYKPMQLQROFITALPARGLDHIAENILSYLDAKSLCAELVCCKEWY 180
181 RVTSDGMLWKLIERVMTDSLWRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
181 RVTSDGMLWKLIERVMTDSLWRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
241 QDIETIESNRWRCGRHSIQRHCRSETSKGVYCLQYDDQKIVSGLRNTIKIWDKNTLECK 300
241 QDIETIESNRWRCGRHSIQRHCRSETSKGVYCLQYDDQKIVSGLRNTIKIWDKNTLECK 300
301 RILTGTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLIIHCEAVLHLRFNNGMM 360
301 RILTGTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLIIHCEAVLHLRFNNGMM 360
361 VTCSDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSASGDRTIKWNTSTC 420
361 VTCSDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSASGDRTIKWNTSTC 420
421 EFVRLNGHKGRIACQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
421 EFVRLNGHKGRIACQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCLRTLVEHSGRVRFLQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCLRTLVEHSGRVRFLQDFEQIVSSSHDDT 540
541 ILIWDFLNDPAAQAEPPSPSRITYYSR 569
541 ILIWDFLNDPAAQAEPPSPSRITYYSR 569

96697 standard; Protein; 569 AA.
96697;
-SEP-2000 (first entry)
nan beta-TrCP.
ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
i-inflammatory; immunosuppressive; cytostatic.
o sapiens.
00034447-A2.
JUN-2000.
DEC-1999; 99WO-US29371.
DEC-1998; 98US-0210060.
GN-) SIGNAL PHARM INC.
SS) YISSUM RES & DEV CO.
ning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
on I, Yaron A;

XX WP; 2000-431294/37.
DR N-ESDB; AAA51229.
XX Polypeptide enhancing phosphorylated I-kappaB ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX Claim 21; Page 72-74; 77pp; English.
XX Human beta-TrCP, an F-box/WD protein family member, has been shown to
CC have homology to human E3 ubiquitin ligase (E3). E3 enhances
CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of
CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
CC degradation via the ubiquitin pathway is useful for identifying
CC modulators of NF-kappa-B. In vitro analysis suggests that deletion of
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
CC the F-box results in a protein that functions as a dominant negative
CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a
CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
XX Sequence 569 AA;
SQ

Query Match 100.0%; Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.5e-286;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLQEKALFNMSSERDCNNGEPKRIIPEKNSLRQTYNSCARCLNQETVCLA 60
Db 1 MDPAAVLQEKALFNMSSERDCNNGEPKRIIPEKNSLRQTYNSCARCLNQETVCLA 60
QY 61 STAMKTENCVAKTTLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
Db 61 STAMKTENCVAKTTLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
QY 121 ISQCHYQHGHSINLYKPMQLQROFITALPARGLDHIAENILSYLDAKSLCAELVCCKEWY 180
Db 121 ISQCHYQHGHSINLYKPMQLQROFITALPARGLDHIAENILSYLDAKSLCAELVCCKEWY 180
QY 181 RVTSDGMLWKLIERVMTDSLWRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
Db 181 RVTSDGMLWKLIERVMTDSLWRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
QY 241 QDIETIESNRWRCGRHSIQRHCRSETSKGVYCLQYDDQKIVSGLRNTIKIWDKNTLECK 300
Db 241 QDIETIESNRWRCGRHSIQRHCRSETSKGVYCLQYDDQKIVSGLRNTIKIWDKNTLECK 300
QY 301 RILTGTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLIIHCEAVLHLRFNNGMM 360
Db 301 RILTGTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLIIHCEAVLHLRFNNGMM 360
QY 361 VTCSDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSASGDRTIKWNTSTC 420
Db 361 VTCSDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSASGDRTIKWNTSTC 420
QY 421 EFVRLNGHKGRIACQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
Db 421 EFVRLNGHKGRIACQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCLRTLVEHSGRVRFLQDFEQIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCLRTLVEHSGRVRFLQDFEQIVSSSHDDT 540
QY 541 ILIWDFLNDPAAQAEPPSPSRITYYSR 569
Db 541 ILIWDFLNDPAAQAEPPSPSRITYYSR 569

4
1
AY83041 standard; Protein: 569 AA.
AY83041;
5-AUG-2000 (first entry)
-box protein FBP-1.
box protein; FBP; diagnosis; treatment; screening; agonist;
agonist; proliferative disorder; differentiative disorder;
cancer; prostate cancer; ovarian cancer; cancer;
cell lung carcinoma; immune disorder; cardiovascular disorder;
inflammatory disorder; human.
mo sapiens.
200012679-A1.
-MAR-2000.
-AUG-1999; 99WO-US19560.
-AUG-1998; 98US-0098355.
-FEB-1999; 99US-0118568.
-MAR-1999; 99US-0124449.
YNY) UNIV NEW YORK STATE.
aur DS, Pagano M, Latres E;
I; 2000-256635/22.
PSDB; AAZ93350.
vel nucleic acid for screening compounds useful for treating
oliferative and differentiative disorders such as cancer and immune
sorders comprises sequences encoding ubiquitin ligases -
closure; Figure 3a; 245pp; English.
leic acids encoding substrate-targeting subunits of ubiquitin
ases with F-box motifs (F-box proteins) are useful for diagnosis
roliferative and differentiated related disorders by measuring
gene expression. Cells expressing such proteins or
ir fragments are useful for screening compounds. The compounds
agonists or antagonists, which are useful for treating a
liferative or differentiative disorder in a mammal such as
ast, ovarian and prostate cancer and small cell lung carcinoma
also major opportunistic infections, immune disorders,
divascular diseases and inflammatory disorders. FBP protein,
logs, derivatives and their subsequences, anti-FBP antibodies
also useful in diagnosis of the disorders.
quence 569 AA;
Match 100.0%; Score 3034; DB 21; Length 569;
ocal Similarity 100.0%; Pred. No. 2.5e-286;
s 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MDPAAVLQKALFPMNSRERDNNCEPPPKIIPEKNSURQYNSCARLCLNQETVCLA 60
1 MDPAAVLQKALFPMNSRERDNNCEPPPKIIPEKNSURQYNSCARLCLNQETVCLA 60
61 STAMKTCNVAKTLANGTSSMIVPKORKLSASYEKELCVKVFQWSESQDVEFVHL 120
61 STAMKTCNVAKTLANGTSSMIVPKORKLSASYEKELCVKVFQWSESQDVEFVHL 120
121 ISQCHYGHGCHNSYLKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
121 ISQCHYGHGCHNSYLKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
181 RVTSDGMLWKLLIERMVRTDSLWGLAERRGWGYLFPKPPDGNAPPNSFYRALYPKII 240

Db 181 RVTSDGMLWKLLIERMVRTDSLWGLAERRGWGYLFPKPPDGNAPPNSFYRALYPKII 240
QY 241 QDIETIESNRCGRHSRLRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
Db 241 QDIETIESNRCGRHSRLRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
QY 301 RLITGHTGSLVCLQYDERVITGSSDSTVRVMDVNTGEMTLTHHCEAVLHFRFNNGM 360
Db 301 RLITGHTGSLVCLQYDERVITGSSDSTVRVMDVNTGEMTLTHHCEAVLHFRFNNGM 360
QY 361 VTCSKDRSTAVWDMASPTDITLRLVGVHRAAVNVDFDDKXIIVSASGDRFIKWNVTSTC 420
Db 361 VTCSKDRSTAVWDMASPTDITLRLVGVHRAAVNVDFDDKXIIVSASGDRFIKWNVTSTC 420
QY 421 EFVRLTNGHKGRTIACLOYRDLRVSGSSDNTIRLWDIECGACLRVLEGHLELVCIRFDN 480
Db 421 EFVRLTNGHKGRTIACLOYRDLRVSGSSDNTIRLWDIECGACLRVLEGHLELVCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTIVEHSGRVFRLQDFEFOIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTIVEHSGRVFRLQDFEFOIVSSSHDDT 540
QY 541 ILIWDPLNDPAAQAEPSPRSTYTYISR 569
Db 541 ILIWDPLNDPAAQAEPSPRSTYTYISR 569
RESULT 5
AAY83250
ID AAY83250 standard; Protein: 569 AA.
AC AAY83250;
XX 16-AUG-2000 (first entry)
DT F-box protein hBetaTrCp.
DE Ubiquitin ligase; SCF; F-box protein; targeted degradation;
XX destabilisation; proteolysis; drug discovery; gene therapy; cancer;
KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
KW human.
XX Homo sapiens.
EN W0200022110-A2.
FD 20-APR-2000.
XX 08-OCT-1999; 99WO-US23705.
XX 09-OCT-1998; 98US-0103787.
XX (HARD) HARVARD COLLEGE.
XX Zhou P, Howley P;
XX WPI; 2000-317970/27.
DR N-PSDB; AAZ933710.
XX Targeting degradation of polypeptide useful for treating cancer and
PT other proliferative disorders, involves conjugating polypeptide with
PT ubiquitin protein ligase or inhibiting ubiquitination using organic
PT compound
XX Claim 9; Page 171; 185pp; English.
XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
CC ligases) which can be used for the targeted degradation of a target
CC polypeptide in vivo. Targeted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or

increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating diseases caused by the presence of normal amount of the specific polypeptides, for drug discovery and for gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

sequence 569 AA;
Match 100.0%; Score 3034; DB 21; Length 569;
Local Similarity 100.0%; Pred No. 2.5e-286;
ies 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MDPAEVLQEKALFPMNSREDCCNGEPKRIPEKNSLRQYNSCARLCLNQETVCLA 60
1 MDPAEVLQEKALFPMNSREDCCNGEPKRIPEKNSLRQYNSCARLCLNQETVCLA 60
61 STAMKTCNVAKTKLANGTSMVVKORKLSASVEKEKLCVYKPFQWSESDQVEVHL 120
61 STAMKTCNVAKTKLANGTSMVVKORKLSASVEKEKLCVYKPFQWSESDQVEVHL 120
121 ISQCHYOHGHINSYLPMLQRODFTALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
121 ISQCHYOHGHINSYLPMLQRODFTALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
181 RVTSDGMLWKKLIERMVRTSLRGLAERGWQYLFKNKPPDGNAPPNSFYRALYPKII 240
181 RVTSDGMLWKKLIERMVRTSLRGLAERGWQYLFKNKPPDGNAPPNSFYRALYPKII 240
241 QDIETISNRCGRHSQRHCRSETSGVYCLQYDDQKIVSGRLONTIKIWDKNTLECK 300
241 QDIETISNRCGRHSQRHCRSETSGVYCLQYDDQKIVSGRLONTIKIWDKNTLECK 300
301 RILTHGTGSLVCLQYDERVITGSSDSTVRVWDVNTGEMTLTIHCEAVLHLRFNNGMM 360
301 RILTHGTGSLVCLQYDERVITGSSDSTVRVWDVNTGEMTLTIHCEAVLHLRFNNGMM 360
361 VTCSKDRSIAVMDWASPTDILRLVLGHRAAVNVDFDKIVSAGDRTIKVWNTSIC 420
361 VTCSKDRSIAVMDWASPTDILRLVLGHRAAVNVDFDKIVSAGDRTIKVWNTSIC 420
421 EFVRTLNGHKGKGIACQYDRDLVVGSSDNTIRLWDIECGACLRVLEGEELVRCIRFDN 480
421 EFVRTLNGHKGKGIACQYDRDLVVGSSDNTIRLWDIECGACLRVLEGEELVRCIRFDN 480
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTICLRTLVEHSGRVFRLQDFEIQVSSHDPT 540
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTICLRTLVEHSGRVFRLQDFEIQVSSHDPT 540
541 ILIWDFLNDPAAQAEPSPSRRTYTIISR 569
541 ILIWDFLNDPAAQAEPSPSRRTYTIISR 569

'44249 standard; Protein; 569 AA.

'44249;

PEB-2000 (first entry)

an cell signalling protein-12.

1 signalling protein-12; CSIGP-12; cell proliferation;
lamatory disorder; cirrhosis; cancer; hepatitis; AIDS;
eriosclerosis; Addison's disease; multiple sclerosis.

OS	Homo sapiens.	
XX	Key	Location/Qualifiers
PH	Modified-site	19
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	39
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	91
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	109
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	162
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	266
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	288
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	328
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	376
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	381
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	411
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	418
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	451
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	514
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	519
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	535
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	536
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	17
FT	Modified-site	/note= "Potential glycosylation site"
FT	Modified-site	77
FT	Modified-site	/note= "Potential glycosylation site"
FT	Modified-site	416
FT	Modified-site	/note= "Potential glycosylation site"
FT	Region	320..334
FT	Region	/label= Signature_sequence
FT	Region	360..374
FT	Region	/label= Signature_sequence
FT	Region	403..417
FT	Region	/label= Signature_sequence
FT	Region	443..457
FT	Region	/label= Signature_sequence
FT	Region	483..497
FT	Region	/label= Signature_sequence
FT	Region	532..546
FT	Region	/label= Signature_sequence
XX	WO9958558-A2.	
PN		
XX		
PD	18-NOV-1999.	
XX		
PF	13-MAY-1999;	99WO-US10567.
XX		
PR	13-MAY-1998;	98US-0085343.
PR	26-AUG-1998;	98US-0098010.
XX		
XX	(INCY-) INCYTE PHARM INC.	
XX	Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;	
PI	Baughn MR, Yang J;	
XX		
XX	WPI; 2000-086432/07.	
DR	N-ESDB; AAZ29233.	
XX		

uman cell signaling proteins useful for, e.g. diagnosing cell
proliferative and inflammatory disorders

laim 1; Page 77-78; 90pp; English.

he present sequence is cell signalling protein-12 (CSIGP-12) encoded
y cDNA obtained from Incyte clone 3239149 of COLAUC701 library. It is
xpressed in musculoskeletal, gastrointestinal and nervous tissues and is
ound to be homologous to beta-transducin repeats containing
rotein. Fragments of CSIGP encoding nucleic acid can be used as
ybridisation probe for detecting CSIGP related sequences or allelic
ariants. Recombinant CSIGP can be produced in host cells by transforming
hem with genetically engineered vectors. Agonists or antagonists can be
sed in the treatment of cell proliferative and inflammatory disorders
associated with decreased or increased CSIGP expression. CSIGP is used in
he diagnosis, prevention and treatment of cell proliferative disorders
ke arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
isorders like AIDS, Addison's disease, multiple sclerosis, etc.

sequence 569 AA;

f Match 100.0%; Score 3034; DB 21; Length 569;
Local Similarity 100.0%; Pred. No. 2.5e-286;
es 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPAEAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

1 MDPAEAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

61 STAMKTCNCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

61 STAMKTCNCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

121 ISOMCHYQHGHSINYLKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

121 ISOMCHYQHGHSINYLKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240

181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240

241 QDIETIESNWRGCRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

241 QDIETIESNWRGCRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

301 RILTGHTGSLVLCQYDERVITGSSDSTVRVWDVNTGEMTLTIHCEAVLHFRNNGWM 360

301 RILTGHTGSLVLCQYDERVITGSSDSTVRVWDVNTGEMTLTIHCEAVLHFRNNGWM 360

361 VTCSDRSIAVWDMASPTDITLRRVLVGHRAAVNVVDFDDKIYVSASGDRTIKWNNTSTC 420

361 VTCSDRSIAVWDMASPTDITLRRVLVGHRAAVNVVDFDDKIYVSASGDRTIKWNNTSTC 420

421 EFVRTLNGHKGIAQLQYDRDLVYSGSDNTIRLWDECGACLRVLEGHEELVRCIRFDN 480

421 EFVRTLNGHKGIAQLQYDRDLVYSGSDNTIRLWDECGACLRVLEGHEELVRCIRFDN 480

481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCRLTLVHSGRVRFLQDFEQIVSSSHDDT 540

481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCRLTLVHSGRVRFLQDFEQIVSSSHDDT 540

541 ILIWLFLNDPRAQAEPSPSRRTYIISR 569

541 ILIWLFLNDPRAQAEPSPSRRTYIISR 569

7

322446 standard; Protein; 569 AA.

322446;

-OCT-2002 (first entry)

XX Human F-box protein FBp1 SEQ ID No 2.

XX Cytostatic; immunomodulator; cardiant; antiinflammatory; antimicrobial;
KW proliferative; differentiative disorder; Skp2; F-box protein; cancer;
KW ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;
KW small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;
KW inflammatory disorder; lymphoma; major opportunistic infection;
KW certain cardiovascular disease; human.

XX Homo sapiens.

XX WO200255665-A2.

XX 18-JUL-2002.

XX 07-JAN-2002; 2002WO-US00311.

XX 05-JAN-2001; 2001US-260179P.

XX (UJNY) UNIV NEW YORK STATE.

XX Pagano M;

XX MPI; 2002-599665/64.

XX N-PSDB; AAL41041.

XX Screening compounds for treating proliferative disorders, e.g. breast
PT cancer or prostate cancer, infections or immune disorders, comprises
PT detecting a change in the activity of Skp2 with either p27 or Cks1 -
XX
PS Disclosure; Fig 3; 246pp; English.

XX The invention relates to screening compounds useful for the treatment of
CC proliferative or differentiative disorders comprising detecting a change
CC in the activity of Skp2 (F-box protein). The method is useful for
CC screening compounds for the treatment of proliferative or differentiative
CC disorders, particularly cancer. These compounds include small molecules,
CC or compounds or derivatives or analogues of the new ubiquitin ligases.
CC The compounds are useful for treating diseases such as cancer (e.g.
CC breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell
CC lung carcinoma or parathyroid adenomas), major opportunistic infections,
CC immune disorders, certain cardiovascular diseases or inflammatory
CC disorders. This sequence represents an F-box protein (FBP) relating to
CC the invention.

XX Sequence 569 AA;

Query Match 100.0%; Score 3034; DB 23; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.5e-286;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

DB 1 MDPAEAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

QY 61 STAMKTCNCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

DB 61 STAMKTCNCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

QY 121 ISOMCHYQHGHSINYLKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

DB 121 ISOMCHYQHGHSINYLKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

QY 181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240

DB 181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240

QY 241 QDIETIESNWRGCRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

DB 241 QDIETIESNWRGCRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

QY 301 RILTGHTGSLVLCQYDERVITGSSDSTVRVWDVNTGEMTLTIHCEAVLHFRNNGWM 360

```
|||||
301 RILTGTGVLCLQYDERVITITSSDSTVRVMDVNTGEMNTLIHCEAVLHFRNNGMM 360
361 VTCSKORSIAVWDMASPTDITLRLVVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
361 VTCSKORSIAVWDMASPTDITLRLVVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
541 ILIWFNDPAAQAEPSPSRRTYTYISR 569
541 ILIWFNDPAAQAEPSPSRRTYTYISR 569
```

8
3

G69473 standard; Protein; 569 AA.

G69473;

-OCT-2002 (first entry)

man bait protein beta-TcP1.

man; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
n-insulin diabetes mellitus; obesity; selected interacting domain;
D; protein-protein interaction map; PIM; anorectic; metabolic disorder.

mo sapiens.

200253726-A2.

-JUL-2002.

-DEC-2001; 2001WO-EP15423.

-JAN-2001; 2001US-259377P.

YER-} HYBRIGENICS.

URS } CENT NAT RECH SCI.

grain P, Marullo S, Jockers R;

I; 2002-583612/62.

5SDB; ABS51009.

vel complex of protein-protein interactions in adipocyte cells for
antifungal compounds that modulate the protein-protein interactions
a useful for treating obesity and metabolic disorders -

aim 1; Page -: 125pp; English.

a invention relates to a complex of protein-protein interactions
forming a protein-protein interaction map, PIM) in adipocyte cells as
fined in the specification, or polynucleotides in adipocytes encoding
r the polypeptides. Also included are a recombinant cell expressing the
eracting polypeptides and a method of selecting a modulating compound
adipocyte cells, by cultivating a recombinant host cell on a selective
thium containing a modulating compound and a reporter gene the
pression of which is toxic for the recombinant host cell which is
nsformed with two vectors, where the first vector comprises a
ynucleotide encoding a first hybrid polypeptide and DNA binding
ain and the second vector comprising a polynucleotide encoding a
ond hybrid polypeptide and an activating domain that activates the
tic reporter gene, when the first and second hybrid polypeptides
eract and selecting the modulating compound which inhibits the
wth of the recombinant host cell (i.e. using the yeast two-hybrid

CC system). The complexes are useful for identifying compounds that modulate
CC the protein-protein interactions and useful for treating obesity and
CC metabolic disorders e.g. non-insulin dependent diabetes mellitus,
CC NIDDM. The compound isolated by the method is useful for treating and
CC preventing obesity or metabolic diseases. The interactions between
CC the proteins of the complex further define a set of selected interacting
CC domains, SIP. The present sequence represents a member of the protein
CC complex of the invention, used as the bait protein in the yeast two-
CC hybrid assay.
CC Note: The present sequence was not displayed in the specification but
CC was obtained from its Genbank entry by the indexer.

XX Sequence 569 AA;

Query Match 100.0%; Score 3034; DB 23; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.5e-286;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPAAVLOEKALKFNMSSEREDCNGEPKRIIEKNSLRQTYNSCARCLNQETVCLA 60
Db 1 MDPAAVLOEKALKFNMSSEREDCNGEPKRIIEKNSLRQTYNSCARCLNQETVCLA 60
Qy 61 STAMKTENCVAKTKLANGTSSMIVPKORLSASYEKEKELCVKYPEQWSESDQVFEHL 120
Db 61 STAMKTENCVAKTKLANGTSSMIVPKORLSASYEKEKELCVKYPEQWSESDQVFEHL 120
Qy 121 ISQWCHYQHGHSYLYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAAEVCKEWY 180
Db 121 ISQWCHYQHGHSYLYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAAEVCKEWY 180
Qy 181 RVTSQGLMKKLIETRMVVRTDSLWRGLAERRGQVLFKNKPPDGNAPPNSFYRALYPKII 240
Db 181 RVTSQGLMKKLIETRMVVRTDSLWRGLAERRGQVLFKNKPPDGNAPPNSFYRALYPKII 240
Qy 241 QDIETTESNWRCCGRHSIQIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
Db 241 QDIETTESNWRCCGRHSIQIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
Qy 301 RILTGTGVLCLQYDERVITITSSDSTVRVMDVNTGEMNTLIHCEAVLHFRNNGMM 360
Db 301 RILTGTGVLCLQYDERVITITSSDSTVRVMDVNTGEMNTLIHCEAVLHFRNNGMM 360
Qy 361 VTCSKORSIAVWDMASPTDITLRLVVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
Db 361 VTCSKORSIAVWDMASPTDITLRLVVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
Qy 421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
Db 421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
Qy 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
Qy 541 ILIWFNDPAAQAEPSPSRRTYTYISR 569
Db 541 ILIWFNDPAAQAEPSPSRRTYTYISR 569

RESULT 9

AAB48298

ID AAB48298 standard; protein; 569 AA.

XX AAB48298;

AC AAB48298;

DT 02-APR-2001 (first entry)

XX Human ZP11 protein.

DE S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;

KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;

KW Bad; Bcl-2; tumour; cytostatic.

XX

omo sapiens.
D200075184-A1.
4-DEC-2000.
5-JUN-2000; 2000WO-US15449.
1-JUN-1999; 99US-0137494.
JYYA) UNIV YALE.
iang H, Tsvetkov IM, Kondo T;
PI; 2001-061703/07.
-PSDB; AAC84610.
odulating polypeptide levels in a cell, diagnosing and treating tumor,
volves altering levels of proteins such as S-phase kinase associated
roteins 1, 2 and cullin/CDC53 proteins -
laim 3; Page 130-132; 162pp; English.
e invention relates to methods of altering the polypeptide levels in a
le, using proteins selected from S-phase kinase associated proteins 1
d 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
illin/ CDC53 family of proteins). The method is useful for altering the
vel of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
ypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
tecting tumors, and in monitoring tumor treatment in a mammal. Agents
at modulate interactions between SKP and target proteins are useful for
eating tumours.
quence 569 AA;
Match 99.8%; Score 3027; DB 22; Length 569;
Local Similarity 99.8%; Pred. No. 1.2e-285;
es 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 MDPAEVLQELKALFPMNSERDCNNGEPKRIPEKNSLRQTVNSCARLCLNQETVCLA 60
1 MDPAEVLQELKALFPMNSERDCNNGEPKRIPEKNSLRQTVNSCARLCLNQETVCLA 60
61 STAMKTECNVAKTKLANGTSWIVPKORKLSASVEKEKELCVKFEQWSESDOVFEHL 120
61 STAMKTECNVAKTKLANGTSWIVPKORKLSASVEKEKELCVKFEQWSESDOVFEHL 120
121 ISQCHYQGHGHSNLYKPMQLQDFITLAPARGLDHIHAEINILSYLDAKSLCAELVCKEWY 180
121 ISQCHYQGHGHSNLYKPMQLQDFITLAPARGLDHIHAEINILSYLDAKSLCAELVCKEWY 180
181 RVTSDGMLWKKLIERWRTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALYPKII 240
181 RVTSDGMLWKKLIERWRTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALYPKII 240
241 QDIETIESNRCGRHSLOIRHCRSETSGKVYCLQYDQDKIVSGLRDNTIKWDKNTLECK 300
241 QDIETIESNRCGRHSLOIRHCRSETSGKVYCLQYDQDKIVSGLRDNTIKWDKNTLECK 300
301 RLITGHTGSLVCLQYDERVVIITGSSDSTVRVWDNTGEMLNTLIHCEAVLHLPFNGWM 360
301 RLITGHTGSLVCLQYDGRVVIITGSSDSTVRVWDNTGEMLNTLIHCEAVLHLPFNGWM 360
361 VTCSKDRSIYVDMASPTDITLRLVYLGHRAAVNVVDFDKYIVSASGDRITIKWNTSTC 420
361 VTCSKDRSIYVDMASPTDITLRLVYLGHRAAVNVVDFDKYIVSASGDRITIKWNTSTC 420
421 EFVRTLNGHKRGIAQLQVDRDLVWVGSGSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
421 EFVRTLNGHKRGIAQLQVDRDLVWVGSGSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKVWDVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKVWDVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
Qy 541 ILIWDFLNDPAQAQAEPRSPRSTYTYISR 569
Db 541 ILIWDFLNDPAQAQAEPRSPRSTYTYISR 569
RESULT 10
AAM00960
ID AAM00960 standard; Protein; 608 AA.
XX
AC AAM00960;
XX
DT 01-OCT-2001 (first entry)
XX
DE Human bone marrow protein, SEQ ID NO: 436.
XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
OS Homo sapiens.
XX WO200153453-A2.
XX 26-JUL-2001.
XX 23-DEC-2000; 2000WO-US34960.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250593.
XX
PA (HYSE-) HYSE INC.
XX
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX
DR WPI; 2001-488707/53.
DR N-PSDB; AAH90079.
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
for treating e.g. cancer and immune deficiency disorders -
XX
PS Claim 10; Page 523-524; 648pp; English.
XX
CC The present sequence is one of 251 novel human polypeptides encoded
by a bone marrow-expressed polynucleotide. The polynucleotide and the
polypeptide encoded by it are useful in the treatment of various
immune deficiencies and disorders. The deficiencies and disorders may
be genetic, may be caused by a viral (e.g. HIV) bacterial or fungal
infection, or may result from an autoimmune disorder, a coagulation
disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
suppression of an inflammatory response or treatment of a nervous
system disorder such as Alzheimer's disease. Detection of the presence
or increased expression of the polynucleotide or the protein it
encodes is useful for the diagnosis and/or prognosis of one
or more types of cancer. The polynucleotide and polypeptide can be
used as nutritional sources or supplements and in the screening of
chemical compounds as potential drugs.
XX
XX Sequence 608 AA;
Query Match 99.8%; Score 3027; DB 22; Length 608;
Best Local Similarity 99.6%; Pred. No. 1.4e-285;

```
tes 567; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 MDPAAVLQEKALFNMSSREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 60
40 MDPAAVLQEKALFNMSSREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 99
61 STAMKTENCVAKTTLANGTSSMIVPKOKLSASVEKEKELCVKVFQWSESDDQVEVEHL 120
100 STAMKTENCVAKTTLANGTSSMIVPKOKLSASVEKEKELCVKVFQWSESDDQVEVEHL 159
121 ISQCHYQGHGHSYLYKPMQLQDPIITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 180
160 ISQCHYQGHGHSYLYKPMQLQDPIITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 219
181 RVSDGMLWKKLIERMVRTDSLWGLAERRGGOYLFPKNPPDGNAPPNSFYRALYPKII 240
220 RVSDGMLWKKLIERMVRTDSLWGLAERRGGOYLFPKNPPDGNAPPNSFYRALYPKII 279
241 QDIETIESNWCGRHSQRHICRSETSGVYCYQYDDQKIVSGLRONTIKIWKNTLECK 300
280 QDIETIESNWCGRHSQRHICRSETSGVYCYQYDDQKIVSGLRONTIKIWKNTLECK 339
301 RILGHTGSVLCQYDERVITGSDSTVRVWVNTGEMNTLIHCEAVLHLRFNNGMM 360
340 RILGHTGSVLCQYDERVITGSDSTVRVWVNTGEMNTLIHCEAVLHLRFNNGMM 399
361 VTCSDRSIAVMDASPTDITLRLVGHRAAVNVVDFDDKIYVSAGDRTIKVWNTSTC 420
400 VTCSDRSIAVMDASPTDITLRLVGHRAAVNVVDFDDKIYVSAGDRTIKVWNTSTC 459
421 EFVTLNGHKGRIACLOVRDLRVVSGSDNTIRLWDIECGACLRVLEHGEELVRCIRFDN 480
460 EFVTLNGHKGRIACLOVRDLRVVSGSDNTIRLWDIECGACLRVLEHGEELVRCIRFDN 519
481 KRIYSGAYDGIKYWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDIT 540
520 KRIYSGAYDGIKYWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDIT 579
541 ILIWDFLNDPAAQEPSPRSRTYIISR 569
580 ILIWDFLNDPAAQEPSPRSRTYIISR 608
11
2
M78582 standard; Protein; 605 AA.
-NOV-2001 (first entry)
man protein SEQ ID NO 1244.
man; cytokine; cell proliferation; cell differentiation; gene therapy;
ccine; peptide therapy; stem cell growth factor; haematopoiesis;
issue growth factor; immunomodulatory; cancer; leukaemia;
rvous system disorder; arthritis; inflammation.
mo sapiens.
200157190-A2.
-AUG-2001.
-FEB-2001; 2001WO-0504098.
-FEB-2000; 2000US-0496914.
-APR-2000; 2000US-0560875.
-JUN-2000; 2000US-0598075.
-JUL-2000; 2000US-0620325.
-SEP-2000; 2000US-0654936.
-SEP-2000; 2000US-0663561.
-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao Qa, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAK51715.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX Claim 20; Page 3503-3504; 6221pp; English.
PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX Sequence 605 AA;
Query Match 99.1%; Score 3006; DB 22; Length 605;
Best Local Similarity 94.0%; Pred. No. 1.5e-283;
Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;
QY 1 MDPAAVLQEKALK-----FMNSSREDC 24
DB 1 MDPAAVLQEKALKFMCMPRLWLGSSLDASMPSLCLYNGPGTALAFNMSSREDC 60
QY 25 NNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTENCVAKTTLANGTSSMIV 84
DB 61 NNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTENCVAKTTLANGTSSMIV 120
QY 85 PKOKLSASVEKEKELCVKVFQWSESDDQVEVEHLISQCHYQGHGHSYLYKPMQLQD 144
DB 121 PKOKLSASVEKEKELCVKVFQWSESDDQVEVEHLISQCHYQGHGHSYLYKPMQLQD 180
QY 145 ITALPARGLDHIAENILSYLDAKSLCAAEVLCKEYRVVTSDDGMLWKKLIERMVRTDSLWR 204
DB 181 ITALPARGLDHIAENILSYLDAKSLCAAEVLCKEYRVVTSDDGMLWKKLIERMVRTDSLWR 240
QY 205 GLAERRGGOYLFPKNPPDGNAPPNSFYRALYPKIIQDIETIESNWCGRHSQRHICRS 264
DB 241 GLAERRGGOYLFPKNPPDGNAPPNSFYRALYPKIIQDIETIESNWCGRHSQRHICRS 300
QY 265 ETSKGVYCYQYDDQKIVSGLRDNTIKIWKNTLECKRILTGHTGSVLCQYDERVITGS 324
DB 301 ETSKGVYCYQYDDQKIVSGLRDNTIKIWKNTLECKRILTGHTGSVLCQYDERVITGS 360
QY 325 SDSTVRVWVNTGEMNTLIHCEAVLHLRFNNGMMVTCSDRSIAVMDASPTDITLRR 384
DB 361 SDSTVRVWVNTGEMNTLIHCEAVLHLRFNNGMMVTCSDRSIAVMDASPTDITLRR 420
QY 385 VLVGHRAAVNVVDPDDKIYVSAGDRTIKVWNTSTCEVFTLNGHKGRIACLOVRDLVV 444
DB 421 VLVGHRAAVNVVDPDDKIYVSAGDRTIKVWNTSTCEVFTLNGHKGRIACLOVRDLVV 480
QY 445 SGSSDNTIRLWDIECGACLRVLEHGEELVRCIRFDNKRIVSGAYDGIKYWDLVAALDPR 504
DB 481 SGSSDNTIRLWDIECGACLRVLEHGEELVRCIRFDNKRIVSGAYDGIKYWDLVAALDPR 540
```

505 APAGTLCRTLVHSGRVFRLQDFEFOIVSSSHDDTILIWDFLNDPAAQAEPPSPRTY 564
 541 APAGTLCRTLVHSGRVFRLQDFEFOIVSSSHDDTILIWDFLNDPAAQAEPPSPRTY 600
 565 TYISR 569
 601 TYISR 605
 12
 12
 AB12812 standard; protein; 569 AA.
 AB12812;
 7-NOV-2000 (first entry)
 mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.
 ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;
 beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
 gene therapy; colon cancer; beta-transducin repeat containing protein;
 beta-TrCP.
 15 musculus.
 2000166542-A.
 1)-JUN-2000.
 -DEC-1998; 98JP-0343437.
 -DEC-1998; 98JP-0343437.
 (AGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 1; 2000-485550/43.
 PSDB; AAA73131.
 box protein of ubiquitin ligase SCF complex which promotes the
 ubiquitination of IkappaB or beta-catenin -
 aim 2; Page 9-10; 19pp; Japanese.
 e present invention describes an F-box motif protein of ubiquitin
 ligase SCF complex which promotes the ubiquitination of IkappaB or
 beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
 motif (SCF complex) of F-box protein containing F-box motif and a
 motif and has the amino acid sequence of 45 residues (AAB12811)
 one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
 ligase FWD1 protein) and (AAB12813, which is human beta-transducin
 repeat containing protein (beta-TrCP)). The F-box protein can be used for
 gene therapy of colon cancer by being recombined to a virus vector.
 quence 569 AA;
 Match
 Local Similarity 98.8%; Score 2997; DB 21; Length 569;
 es 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 1 MDPAAVLQEKALKFNNSSREDCCNGGPPPKIIPKNSLRQTYNSCARLCLNQETVCLA 60
 1 MDPAAVLQEKALKFNNSSREDCCNGGPPPKIIPKNSLRQTYNSCARLCLNQETVCLT 60
 61 STAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKELCVKVFQWSESDDQVFVEHL 120
 61 STAMKTENCVAKALANGTSSMIVPKQKLSASYEKELCVKVFQWSESDDQVFVEHL 120
 121 ISQCHYQGHGINSYKPLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
 121 ISQCHYQGHGINSYKPLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
 181 RVTSQGLWKKLIERMVRTDSLWRGLAERRGWSQYLFKNKPPDENAPPNSFYRALYPKII 240

DB 181 RVTSQGLWKKLIERMVRTDSLWRGLAERRGWSQYLFKNKPPDENAPPNSFYRALYPKII 240
 QY 241 ODIEETISNWRGHSIORIHCSETSKGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
 DB 241 ODIEETISNWRGHSIORIHCSETSKGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
 QY 301 RILTGHTGSLVCLQYDERVIIITGSSDSTVRVNDVNTGEMLNTLIHCEAVLHRLFNNGMM 360
 DB 301 RILTGHTGSLVCLQYDERVIIITGSSDSTVRVNDVNTGEMLNTLIHCEAVLHRLFNNGMM 360
 QY 361 VTCSDRSIAVWDMASPTDITLRLVLVGHRAAANNVDFDDKYIVSASGDRTIKWNNTSTC 420
 DB 361 VTCSDRSIAVWDMASPTDITLRLVLVGHRAAANNVDFDDKYIVSASGDRTIKWNNTSTC 420
 QY 421 EFVRLNGHKGACIQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
 DB 421 EFVRLNGHKGACIQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
 QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVLHSGRVFRLQDFEFOIVSSSHDDT 540
 DB 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVLHSGRVFRLQDFEFOIVSSSHDDT 540
 QY 541 ILIWDFLNDPAAQAEPPSPRTYTYISR 569
 DB 541 ILIWDFLNDPAAQAEPPSPRTYTYISR 569
 RESULT 13
 AAY83254
 ID AAY83254 standard; Protein; 569 AA.
 AC AAY83254;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein FWD1p.
 XX
 KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 KW destabilization; proteolysis; drug discovery; gene therapy; cancer;
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
 KW mouse; ss.
 OS
 XX Mus musculus.
 XX
 PN WO200022110-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US23705.
 XX
 PR 09-OCT-1998; 98US-0103787.
 XX
 BA (HARD) HARVARD COLLEGE.
 XX
 PI Zhou P, Howley P;
 XX
 DR WPI; 2000-317970/27.
 DR N-PSDB; AAY93714.
 XX
 PT Targeting degradation of polypeptide useful for treating cancer and
 PT other proliferative disorders, involves conjugating polypeptide with
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 compound
 XX
 PS Claim 9; Page 184-185; 185pp; English.
 XX
 CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 CC ligases) which can be used for the targeted degradation of a target
 CC polypeptide in vivo. Targeted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or

increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating diseases caused by the presence of abnormal amount of the specific polypeptides, for drug discovery and or gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

sequence 569 AA;

Y Match 98.8%; Score 2997; DB 21; Length 569;

Local Similarity 98.6%; Pred. No. 1e-282;

hes 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

1 MDPAAVLQEKALFPMNSEREDCNGEPPPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

1 MDPAAVLQEKALFPMNSEREDCNGEPPPKIIPEKNSLRQTYNSCARLCLNQETVCLT 60

61 STAMKTCNCVAKTKLANGTSMTVPKORKLSASVEKEKELCVKYFEOWSSDQVEFVHL 120

61 STAMKTCNCVAKTKLANGTSMTVPKORKLSASVEKEKELCVKYFEOWSSDQVEFVHL 120

121 ISQCHYQGHINSYLPKMLQDFITLALPARGLDHIAENILSYLDAKSLCAAEVLCKEWY 180

121 ISQCHYQGHINSYLPKMLQDFITLALPARGLDHIAENILSYLDAKSLCAAEVLCKEWY 180

181 RVTSQMLWKKLIERNVRTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALPKII 240

181 RVTSQMLWKKLIERNVRTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALPKII 240

241 QDIETIESNWRGCRHSLOIHCRTSETSGVYCLQYDDQKIVSGLRDNITIKWDKNTLECK 300

241 QDIETIESNWRGCRHSLOIHCRTSETSGVYCLQYDDQKIVSGLRDNITIKWDKNTLECK 300

301 RILTGHTGSVLQYDERVITIGSSDSTVRVWVNTGEMNTLIHHCCEAVLHLRFNNGWM 360

301 RILTGHTGSVLQYDERVITIGSSDSTVRVWVNTGEMNTLIHHCCEAVLHLRFNNGWM 360

361 VTCKORSIAVNDWASPTDITLRVLVGHRAAVNVDFDKIVYSAGSDTIKWNVTSTC 420

361 VTCKORSIAVNDWASPTDITLRVLVGHRAAVNVDFDKIVYSAGSDTIKWNVTSTC 420

421 EFVRTLNGHKGRIACLOVYRDLVWVGSSDNTIRLWDIECGACLRVLEGEELVRCIRFDN 480

421 EFVRTLNGHKGRIACLOVYRDLVWVGSSDNTIRLWDIECGACLRVLEGEELVRCIRFDN 480

481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540

481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540

541 ILIIMDFLNDPAAQEPSPSRRTYISR 569

541 ILIIMDFLNDPAAQEPSPSRRTYISR 569

M78584 standard; Protein; 632 AA.

M78584;

-NOV-2001 (first entry)

nan protein SEQ ID NO 1246.

nan; cytokine; cell proliferation; cell differentiation; gene therapy;

scine; peptide therapy; stem cell growth factor; haematopoiesis;

seue growth factor; immunomodulatory; cancer; leukaemia;

rvious system disorder; arthritis; inflammation.

XX OS Homo sapiens.

XX FN WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663541.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX DR WPI; 2001-476283/51.

XX DR N-ESDB; AAK51717.

XX Nucleic acids encoding polypeptides with cytokine-like activities,

XX useful in diagnosis and gene therapy -

XX PS Claim 20; Page 3505-3507; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

XX SQ Sequence 632 AA;

Query Match 98.6%; Score 2992.5; DB 22; Length 632;

Best Local Similarity 90.0%; Pred. No. 3.3e-282;

Matches 569; Conservative 0; Mismatches 0; Indels 63; Gaps 1;

QY 1 MDPAAVLQEKALX-----FMNSEREDCNGEPPPKIIPEKNSLRQTYNSCARLCLNQETV 14

DB 1 MDPAAVLQEKALXFMNSEREDCNGEPPPKIIPEKNSLRQTYNSCARLCLNQETV 60

QY 15 -----FMNSEREDCNGEPPPKIIPEKNSLRQTYNSCARLCLNQETV 57

DB 61 MPELRCLYNPGTGALTAFMNSEREDCNGEPPPKIIPEKNSLRQTYNSCARLCLNQETV 120

QY 58 CLASTAMKTCNCVAKTKLANGTSMTVPKORKLSASVEKEKELCVKYFEOWSSDQVEFV 117

DB 121 CLASTAMKTCNCVAKTKLANGTSMTVPKORKLSASVEKEKELCVKYFEOWSSDQVEFV 180

QY 118 EHLISQCHYQGHINSYLPKMLQDFITLALPARGLDHIAENILSYLDAKSLCAAEVLCK 177

DB 181 EHLISQCHYQGHINSYLPKMLQDFITLALPARGLDHIAENILSYLDAKSLCAAEVLCK 240

QY 178 EMYVTSQMLWKKLIERNVRTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALYP 237

DB 241 EMYVTSQMLWKKLIERNVRTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALYP 300

238 KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL 297
 |||||
 301 KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL 360
 |||||
 298 ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWVNTGEMLNTLIHHCCEAVLHLPFNN 357
 |||||
 361 ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWVNTGEMLNTLIHHCCEAVLHLPFNN 420
 |||||
 358 GMMVTCSDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIIVSASGDRTIKWNT 417
 |||||
 421 GMMVTCSDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIIVSASGDRTIKWNT 480
 |||||
 418 STCEFVRTLNHGKRGIAQLQYDERLVVSGSSDNTIRLWDIECGACLRVLEHGBELVRCIR 477
 |||||
 481 STCEFVRTLNHGKRGIAQLQYDERLVVSGSSDNTIRLWDIECGACLRVLEHGBELVRCIR 540
 |||||
 478 FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFLQDFEFOIVSSSH 537
 |||||
 541 FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFLQDFEFOIVSSSH 600
 |||||
 538 DDTILWDFLNDPAAQAEPSPSRSTYTIYSR 569
 |||||
 601 DDTILWDFLNDPAAQAEPSPSRSTYTIYSR 632
 |||||

15
 56

AM79566 standard; Protein; 654 AA.

AM79566;

5-NOV-2001 (first entry)

uman protein SEQ ID NO 3212.

uman; cytokine; cell proliferation; cell differentiation; gene therapy;
 acine; peptide therapy; stem cell growth factor; haematopoiesis;
 issue growth factor; immunomodulatory; cancer; leukaemia;
 rous system disorder; arthritis; inflammation.

mo sapiens.

Y200157190-A2.

-AUG-2001.

-FEB-2001; 2001WO-US04098.

-FEB-2000; 2000US-0496914.

-APR-2000; 2000US-0560875.

-JUN-2000; 2000US-0598075.

-JUL-2000; 2000US-0620325.

-SEP-2000; 2000US-0654936.

-SEP-2000; 2000US-0663561.

-OCT-2000; 2000US-0693325.

-NOV-2000; 2000US-0728422.

I; 2001-476283/51.

PSDB; AAK52699.

clic acids encoding polypeptides with cytokine-like activities,
 eful in diagnosis and gene therapy -

aim 20; Page 285-286; 6221pp; English.

e invention relates to polynucleotides (AAK51456-AAK53435) and the
 coded polypeptides (AAM78323-AAM80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX SQ Sequence 654 AA;

Query Match	98.5%;	Score 2989.5;	DB 22;	Length 654;
Best Local Similarity	89.9%;	Pred. No. 6.9e-282;		
Matches 568;	Conservative 1;	Mismatches 0;	Indels 63;	Gaps 1;
QY	1	MDPAEAVLQEKALFKM-----	-----	16
DB	23	MDPAEAVLQEKALFKMFEPRSWCPGNTMAERSLTATSTSRVQCSMPRSLWLGCSLSADS	82	
QY	17	-----NSSREDNCNGEPPRKIIPEKNSLRQTVNSCARLCLNQETV	57	
DB	83	MPSLACLVPNGTGALTAFONSSEREDNCNGEPPRKIIPEKNSLRQTVNSCARLCLNQETV	142	
QY	58	CLASTAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYFEQWSESDDVEFY	117	
DB	143	CLASTAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYFEQWSESDDVEFY	202	
QY	118	EHLISOMCHYOQHINSYLVKPMLODFITALPARGLDHIAENTILSYLDAKSLCAAEVLCK	177	
DB	203	EHLISOMCHYOQHINSYLVKPMLODFITALPARGLDHIAENTILSYLDAKSLCAAEVLCK	262	
QY	178	EWYRVTSQDGLMKKLIERMVTRTSLMRGLAERGGWQYLFKNKPPDGNAPPNSFYRALYP	237	
DB	263	EWYRVTSQDGLMKKLIERMVTRTSLMRGLAERGGWQYLFKNKPPDGNAPPNSFYRALYP	322	
QY	238	KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL	297	
DB	323	KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL	382	
QY	298	ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWVNTGEMLNTLIHHCCEAVLHLPFNN	357	
DB	383	ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWVNTGEMLNTLIHHCCEAVLHLPFNN	442	
QY	358	GMMVTCSDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIIVSASGDRTIKWNT	417	
DB	443	GMMVTCSDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIIVSASGDRTIKWNT	502	
QY	418	STCEFVRTLNHGKRGIAQLQYDERLVVSGSSDNTIRLWDIECGACLRVLEHGBELVRCIR	477	
DB	503	STCEFVRTLNHGKRGIAQLQYDERLVVSGSSDNTIRLWDIECGACLRVLEHGBELVRCIR	562	
QY	478	FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFLQDFEFOIVSSSH	537	
DB	563	FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFLQDFEFOIVSSSH	622	
QY	538	DDTILWDFLNDPAAQAEPSPSRSTYTIYSR	569	
DB	623	DDTILWDFLNDPAAQAEPSPSRSTYTIYSR	654	

Search completed: October 22, 2003, 10:22:48

Job time : 88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

tein - protein search, using sw model

: October 22, 2003, 10:25:22 ; Search time 72 Seconds
(without alignments)
1323.404 Million cell updates/sec

US-09-601-168B-2
t score: 3034
ce: 1 MDPAAVLQKALKFMSSE.....PAAQEPSPSRITYISR 569

g table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ed: 629382 seqs, 167460630 residues

number of hits satisfying chosen parameters: 629382

m DB seq length: 0

m DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

se : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB ID	Description
3034	100.0	569	12	US-09-832-161-18
3034	100.0	569	14	US-10-042-417-2
3034	100.0	569	15	US-10-038-010-8
2384.5	78.6	542	12	US-09-832-161-16
1163	38.3	219	15	US-10-023-530-2
989.5	32.6	265	9	US-09-764-848-30
989.5	32.6	265	12	US-10-222-020-30
989.5	32.6	265	15	US-10-116-016-30
644.5	21.2	626	10	US-09-213-888-21
644.5	21.2	626	10	US-09-328-877A-21
640	21.1	540	10	US-09-213-888-7
640	21.1	540	10	US-09-213-888-10
640	21.1	540	10	US-09-328-877A-7
640	21.1	540	10	US-09-328-877A-10
640	21.1	540	12	US-10-245-618-14

Sequence 6, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 18, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 6, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 25, Appli
Sequence 25, Appli
Sequence 2, Appli
Sequence 314, Appl
Sequence 3506, Ap
Sequence 8506, Ap
Sequence 30, Appl
Sequence 7658, Ap
Sequence 10, Appl
Sequence 29, Appl
Sequence 8, Appli
Sequence 12, Appli

ALIGNMENTS

RESULT 1

US-09-832-161-18
; Sequence 18, Application US/09832161
; Publication NO. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercurio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinon
; APPLICANT: Davis, Matti
; APPLICANT: Hatzuba, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: NP-KB
; FILE REFERENCE: 860098.427
; CURRENT APPLICATION NUMBER: US/09/832,161
; CURRENT FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: US/98/210,060
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 18
; LENGTH: 569
; TYPE: PPT
; ORGANISM: Homo sapiens
US-09-832-161-18

Query Match 100.0%; Score 3034; DB 12; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.5e-290;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPAAVLQKALKFMSSEDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60

Db 1 MDPAAVLQKALKFMSSEDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60

Qy 61 STAMTENCVAKTLKANGTSSMIVPKOKLSASKEKELCVKYPFQMSSEDOVEFVHL 120

PCB DATE Sept 4 2003

PCB DATE 12-10-98

61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
121 ISQCHYQGHCHINSYKPKMLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
121 ISQCHYQGHCHINSYKPKMLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
181 RVTSDGMLWKXJIERMVRTDSLWRGLAERRGGQYLFKNKPPDGNAPPNSFYRALYPKII 240
181 RVTSDGMLWKXJIERMVRTDSLWRGLAERRGGQYLFKNKPPDGNAPPNSFYRALYPKII 240
241 ODITETESNWRGCHSLQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
241 ODITETESNWRGCHSLQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
301 RILTGTGSLVCLQYDERVIIITGSSDSTVRVMDVNTGEMLNTLIHHCEAVLHLRFNNGMM 360
301 RILTGTGSLVCLQYDERVIIITGSSDSTVRVMDVNTGEMLNTLIHHCEAVLHLRFNNGMM 360
361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDPDDKXIIVSASGDRTIKWNNTSTC 420
361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDPDDKXIIVSASGDRTIKWNNTSTC 420
421 EFVRTLNHKGRIACLOQRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
421 EFVRTLNHKGRIACLOQRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
541 ILIWDFLNDPAAQAEPSPRSRTYTYISR 569
541 ILIWDFLNDPAAQAEPSPRSRTYTYISR 569

042-417-2
ence 2, Application US/10042417
ication No. US20020123082A1
RAL INFORMATION:
LICANT: Pagano, M.
JE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
JE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
3 REFERENCE: 5914-030-999
ENT APPLICATION NUMBER: US/10/042,417
ENT FILING DATE: 2002-01-07
OR APPLICATION NUMBER: 60/260,179
OR FILING DATE: 2001-01-5
ER OF SEQ ID NOS: 89
WARE: Patent In Ver. 2.0
ID NO 2
GTH: 569
E: PRT
ANISM: Homo sapiens
042-417-2
Match 100.0%; Score 3034; DB 14; Length 569;
Local Similarity 100.0%; Pred. No. 1.5e-290;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MDPAAVLOEKALKFNMSSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
1 MDPAAVLOEKALKFNMSSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
121 ISQCHYQGHCHINSYKPKMLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
121 ISQCHYQGHCHINSYKPKMLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
181 RVTSDGMLWKXJIERMVRTDSLWRGLAERRGGQYLFKNKPPDGNAPPNSFYRALYPKII 240

Db 181 RVTSDGMLWKXJIERMVRTDSLWRGLAERRGGQYLFKNKPPDGNAPPNSFYRALYPKII 240
Qy 241 ODITETESNWRGCHSLQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
Db 241 ODITETESNWRGCHSLQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
Qy 301 RILTGTGSLVCLQYDERVIIITGSSDSTVRVMDVNTGEMLNTLIHHCEAVLHLRFNNGMM 360
Db 301 RILTGTGSLVCLQYDERVIIITGSSDSTVRVMDVNTGEMLNTLIHHCEAVLHLRFNNGMM 360
Qy 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDPDDKXIIVSASGDRTIKWNNTSTC 420
Db 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDPDDKXIIVSASGDRTIKWNNTSTC 420
Qy 421 EFVRTLNHKGRIACLOQRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
Db 421 EFVRTLNHKGRIACLOQRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
Qy 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
Qy 541 ILIWDFLNDPAAQAEPSPRSRTYTYISR 569
Db 541 ILIWDFLNDPAAQAEPSPRSRTYTYISR 569

RESULT 3
US-10-038-010-8
; Sequence 8, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 8
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: beta-TrCP1
; LOCATION: (1)..(569)
; OTHER INFORMATION:
US-10-038-010-8

Query Match 100.0%; Score 3034; DB 15; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.5e-290;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPAAVLOEKALKFNMSSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
Db 1 MDPAAVLOEKALKFNMSSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
Qy 61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
Db 61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
Qy 121 ISQCHYQGHCHINSYKPKMLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
Db 121 ISQCHYQGHCHINSYKPKMLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
Qy 181 RVTSDGMLWKXJIERMVRTDSLWRGLAERRGGQYLFKNKPPDGNAPPNSFYRALYPKII 240
Db 181 RVTSDGMLWKXJIERMVRTDSLWRGLAERRGGQYLFKNKPPDGNAPPNSFYRALYPKII 240

241 QDIETIESNRCGRHSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
241 QDIETIESNRCGRHSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
301 RILTHTGTSVLCLOYDERVITITGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNNGM 360
301 RILTHTGTSVLCLOYDERVITITGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNNGM 360
361 VTCSKDRSIAVMDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRITIKVWNTSTC 420
361 VTCSKDRSIAVMDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRITIKVWNTSTC 420
421 EFVRTLNGHKGRIACIQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
421 EFVRTLNGHKGRIACIQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTCLRTLVHSGRVFRLQDFEQIVSSSHDPT 540
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTCLRTLVHSGRVFRLQDFEQIVSSSHDPT 540
541 ILIWDFLNDPAAQAEPSPSRRTTYISR 569
541 ILIWDFLNDPAAQAEPSPSRRTTYISR 569

32-161-16
nce 16, Application US/09832161
cation No. US20030166587A1
AL INFORMATION:
ICANT: Manning, Anthony M.
ICANT: Mercurio, Frank
ICANT: Amit, Sharon
ICANT: Ben-Neriah, Yinon
ICANT: Davis, Metti
ICANT: Hatzubai, Ada
ICANT: Lavon, Iris
ICANT: Yaron, Avraham
E OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
E OF INVENTION: NF-KB
REFERENCE: 860098.427
ENT APPLICATION NUMBER: US/09/832,161
ENT FILING DATE: 2001-04-09
2 APPLICATION NUMBER: 09/210,060
2 FILING DATE: 1998-12-10
ER OF SEQ ID NOS: 30
ARE: Patent in Ver. 2.0
3 NO 16
3TH: 542
5: PRT
ANISM: Homo sapiens
32-161-16

Match 78.6%; Score 2384.5; DB 12; Length 542;
Local Similarity 79.0%; Pred. No. 2e-226;
as 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;
1 MDPAAVLQKALKFNNSSEREDCNNGBPBKIIPEKNSLRQTYNSCARCLNQETVCLA 60
1 MEP-DSVIEDKTIELMCS-----VP-----RSLWLGCANLV---ESMCAL 36
61 S--TAMKTENCVAKTCLANGTSSMIVPKORKLSASYEKEKELCVKYPEOWSESQVFEVHL 118
37 SCLOMPSVRLCL---QISNQTSSIVSRKKPSSGNTQKEDLCIKYFDQWSSDQVEFVE 93
119 HLISQCHYQHGHINSYLPKMLQDFITALPARGLDHIAENILSYLDAKSLCAAEVLCKE 178
94 HLISRMCHYQHGHINSYLPKMLQDFITALPEQGLDHAENILSYLDAKSLCAAEVLCKE 153
179 WYRVSDGMLWKKLIERMVTRDSLWGLAERRGQYLPKPKPDGNAPNPFYALYPK 238
154 WQVVISGMLWKKLIERMVTRDSLWGLAERRGQYLPKPKRPTDG--PNSFYRSLYPK 211

Qy 239 IIODIETIESNRCGRHSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLE 298
Db 212 IIODIETIESNRCGRHSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLE 271
Qy 299 CKRILTHTGTSVLCLOYDERVITITGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNNG 358
Db 272 CLKVLHTGTSVLCLOYDERVITITGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNNG 331
Qy 359 MMVTCSKDRSIAVMDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRITIKVWNTS 418
Db 332 LMVTCSKDRSIAVMDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRITIKVWNTS 391
Qy 419 TCFVRTLNGHKGRIACIQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRF 478
Db 392 TCFVRTLNGHKGRIACIQYDRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRF 451
Qy 479 DNKRIVSGAYDGKIKVMDLVAALDPRAPAGTCLRTLVHSGRVFRLQDFEQIVSSSHD 538
Db 452 DNKRIVSGAYDGKIKVMDLQAALDPRAPAGTCLRTLVHSGRVFRLQDFEQIVSSSHD 511
Qy 539 DTLIWDFLNDPAAQAEPSPSRRTTYISR 569
Db 512 DTLIWDFLNDPAAQAEPSPSRRTTYISR 542

RESULT 5
US-10-023-530-2
; Sequence 2, Application US/10023530
; Publication No. US20030007956A1
; GENERAL INFORMATION:
; APPLICANT: LEGRAIN, Pierre
; APPLICANT: BENAROUS, Richard
; APPLICANT: BLOT, Guillaume
; APPLICANT: LASSOT, Irina
; TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP
; FILE REFERENCE: B4717A
; CURRENT APPLICATION NUMBER: US/10/023,530
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 60/256,276
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: BetaTrCP
; LOCATION: (1)..(219)
; OTHER INFORMATION: F-box protein
US-10-023-530-2

Query Match 38.3%; Score 1163; DB 15; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.2e-106;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPAAVLQKALKFNNSSEREDCNNGBPBKIIPEKNSLRQTYNSCARCLNQETVCLA 60
Db 1 MDPAAVLQKALKFNNSSEREDCNNGBPBKIIPEKNSLRQTYNSCARCLNQETVCLA 60
Qy 61 STAMKTENCVAKTCLANGTSSMIVPKORKLSASYEKEKELCVKYPEOWSESQVFEVHL 120
Db 61 STAMKTENCVAKTCLANGTSSMIVPKORKLSASYEKEKELCVKYPEOWSESQVFEVHL 120
Qy 121 ISQWCHYQHGHINSYLPKMLQDFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 180
Db 121 ISQWCHYQHGHINSYLPKMLQDFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 180
Qy 181 RVTS DGMWKKLIERMVTRDSLWGLAERRGQYLPK 219
Db 181 RVTS DGMWKKLIERMVTRDSLWGLAERRGQYLPK 219

6
764-848-30
ance 30, Application US/09764848
ic No. US2002007270A1
RAL INFORMATION:
ICANT: Rosen et al.
E OF INVENTION: Nucleic Acids, Proteins, and Antibodies
REFERENCE: PT208
ENT APPLICATION NUMBER: US/09764,848
ENT FILING DATE: 2001-01-17
xx application data removed - consult PALM or file wrapper
SER OF SEQ ID NOS: 53
ID NO 30
WARE: Patentin Ver. 2.0
E: PRT
ANISM: Homo sapiens
764-848-30
Match 32.6%; Score 989.5; DB 9; Length 265;
Local Similarity 68.9%; Pred.No. 4.1e-89;
ies 186; Conservative 24; Mismatches 21; Indels 39; Gaps 3;
15 FMNSEREDCNGGPPKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTCVAKTK 74
20 FNTSVMBEDQNEDESPK-----KNTLWQ----- 42
75 LANGTSSMIVPKQKLSASVEKEKELCVKYPEOWSESQDQVEFVHLLISOMCHYQGHINS 134
43 ISNGTSSVIVSRKPSSEGVQKDKICKYFDQWSESQDQVEFVHLLISRMCHYQGHINS 102
135 YKPMQLQBDFTALPARGLDHIAENILSYLDASLCAELVCKEWRVTSQGLMWKLLIE 194
103 YKPMQLQBDFTALPARGLDHIAENILSYLDASLCAELVCKEWRVTSQGLMWKLLIE 162
195 RMVTRDSLWRGLAERGGQVLFKNKPPDGNAPNSFYRALYKPIQDIETIESNWRGR 254
163 RMVTRDPLWGLSERGGQVLFKNRPTDG--PNSFYRSYKPIQDIETIESNWRGR 220
255 HSLQRHCRSETSGVYCLQYDDQKIVSGL 284
221 HNLQRIQCRSENSGVYCLQYDDQKIVSGL 250
7
22-020-30
nce 30, Application US/10222020
cation No. US2003017539A1
AL INFORMATION:
ICANT: Rosen et al.
E OF INVENTION: Nucleic Acids, Proteins, and Antibodies
REFERENCE: PT208C2
ENT APPLICATION NUMBER: US/10/222,020
ENT FILING DATE: 2002-08-16
R APPLICATION NUMBER: 10/116,016
R FILING DATE: 2002-04-05
R APPLICATION NUMBER: 09/764,848
R FILING DATE: 2001-01-17
R APPLICATION NUMBER: 60/179,065
R FILING DATE: 2000-01-31
R APPLICATION NUMBER: 60/180,628
R FILING DATE: 2000-02-04
R APPLICATION NUMBER: 60/214,886
R FILING DATE: 2000-06-28
R APPLICATION NUMBER: 60/217,487
R FILING DATE: 2000-07-11
R APPLICATION NUMBER: 60/225,758
R FILING DATE: 2000-08-14
R APPLICATION NUMBER: 60/220,963
R FILING DATE: 2000-07-26
R APPLICATION NUMBER: 60/217,496
R FILING DATE: 2000-07-11
R APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02

R APPLICATION NUMBER: 60/236,370
R FILING DATE: 2000-09-29
R APPLICATION NUMBER: 60/236,802
R FILING DATE: 2000-10-02
R APPLICATION NUMBER: 60/237,037
R FILING DATE: 2000-10-02
R APPLICATION NUMBER: 60/237,040
R FILING DATE: 2000-10-02
R APPLICATION NUMBER: 60/240,960
R FILING DATE: 2000-10-20
R APPLICATION NUMBER: 60/239,935
R FILING DATE: 2000-10-13
R APPLICATION NUMBER: 60/239,937
R FILING DATE: 2000-10-13
R APPLICATION NUMBER: 60/241,787
R FILING DATE: 2000-10-20
R APPLICATION NUMBER: 60/246,474
R FILING DATE: 2000-11-08
R APPLICATION NUMBER: 60/246,532
R FILING DATE: 2000-11-08
R APPLICATION NUMBER: 60/249,216
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,210
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/226,681
R FILING DATE: 2000-08-22
R APPLICATION NUMBER: 60/225,759
R FILING DATE: 2000-08-14
R APPLICATION NUMBER: 60/225,213
R FILING DATE: 2000-08-14
R APPLICATION NUMBER: 60/227,182
R FILING DATE: 2000-08-22
R APPLICATION NUMBER: 60/225,214
R FILING DATE: 2000-08-14
R APPLICATION NUMBER: 60/235,836
R FILING DATE: 2000-09-27
R APPLICATION NUMBER: 60/230,438
R FILING DATE: 2000-09-06
R APPLICATION NUMBER: 60/215,135
R FILING DATE: 2000-06-30
R APPLICATION NUMBER: 60/225,266
R FILING DATE: 2000-08-14
R APPLICATION NUMBER: 60/249,218
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,208
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,213
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,244
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,212
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,207
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,245
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,244
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,217
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,211
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,215
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,264
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,214
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,297
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/232,400
R FILING DATE: 2000-09-14
R APPLICATION NUMBER: 60/231,242
R FILING DATE: 2000-09-08
R APPLICATION NUMBER: 60/232,081

PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08

Query Match 32.6%; Score 989.5; DB 12; Length 265;
Best Local Similarity 68.9%; Pred. No. 4.1e-89;
Matches 186; Conservative 24; Mismatches 21; Indels 39; Gaps 3;

QY 15 FMSSEREDCNNGEPKRIIPEKNLSRQTYNSCARLCLNQTCLASTAMKTENCVAKTK 74
DB 20 FQNTSVMEDQNEDESPK-----KNTLWQ----- 42
QY 75 LANGTSSMIVPKQKLSASVEKEKELCVKVFQWSESDQVEFVEHLISQCHYQHGHNS 134
DB 43 ISNGTSSVIVSRKRPSEGNQKEXDLCKIFYDQWSESDQVEFVEHLISRMCHYQHGHNS 102
QY 135 YLKPMLQORDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWYRVTSQGLMVKKLE 194
DB 103 YLKPMLQORDFITALPEQGLDHIENILSYLDAKSLCAELVCKEWQVISEGMLVKKLE 162
QY 195 RMVRTDSLWRGLAERQGWQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETIESNWRCCR 254
DB 163 RMVRTDPLWKGLSERRGWDQYLFKNRPTDG--PPNSFYRSLYPKIIQDIETIESNWRCCR 220
QY 255 HSLQRIHCRSETSKGVYCLQYDQKIVSGL 284
DB 221 HNLQRIQCRSENSKGVYCLQYDDEKIISGL 250

RESULT 8
US-10-116-016-30
Sequence 30, Application US/10116016
Publication No. US20030054379A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ08C1
CURRENT APPLICATION NUMBER: US/10/116,016
CURRENT FILING DATE: 2002-04-05
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 30
LENGTH: 265
TYPE: PRI
ORGANISM: Homo sapiens
US-10-116-016-30

Query Match 32.6%; Score 989.5; DB 15; Length 265;

Local Similarity 68.9%; Pred. No. 4.1e-89;
hes 186; Conservative 24; Mismatches 21; Indels 39; Gaps 3;
15 PMNSEREDCNGGPPPKIPEKSLRQTYNSCARLCLNQETVCLASTAMKTCNCVAKYK 74
20 PONTSMEDQNEDESPK-----KNTLWQ-----42
75 LANGTSSMIVPKQKLSASYEKEKELCVKYFEQWSESDQVEFVEHLISQCHYQHGHS 134
43 ISNGTSSVIVSRKPSGEGNYKEDLCKYFDQWSESDQVEFVEHLISRCHYQHGHS 102
135 YLKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELVCKEYRWVTSGMLWKXLE 194
103 YLKPMQLQDFTALPEQGLDHAENILSYLDAKSLCAELVCKEYRWVTSGMLWKXLE 162
195 RWRVTDLSWRGLABRRGWQYLFKNKPPDGNAPNSFVRALYPKIIODIETIESNWRGR 254
163 RWRVTDPLWKLSERRGWQYLFKNRPTDG--PNSFVRSLYPKIIODIETIESNWRGR 220
255 HSLQRIHCRSETSGVYCLQYDDQKIVSGL 284
221 HNLQRIQCRSESGVYCLQYDDQKIVSGL 250
9
213-888-21
nce 21, Application US/09213888A
nt No. US20020164683A1
AL INFORMATION:
JICANT: Gurney, Mark E.
JICANT: Li, Jinhe
JICANT: Pauley, Adele M.
JICANT: Pharmacia & Upjohn Company
E OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
E OF INVENTION: Encode Them
3 REFERENCE: 6142
RENT APPLICATION NUMBER: US/09/213.888A
RENT FILING DATE: 1998-12-17
ER OF SEQ ID NOS: 27
WARE: PatentIn Ver. 2.0
D NO 21
GTH: 626
E: PRT
ANISM: Artificial Sequence
TURE:
IER INFORMATION: Description of Artificial Sequence: 6 myc tagged
IER INFORMATION: homo sapien
213-888-21
Match 21.2%; Score 644.5; DB 10; Length 626;
Local Similarity 30.1%; Pred. No. 1.8e-54;
hes 169; Conservative 103; Mismatches 22; Indels 67; Gaps 15;
8 LOEKALKFMSEREDCNGGPPPKIPEK--NSLRQ--TYNSCARLCLNQETVCLASTA 63
53 MEQKLISEEDLNEMESLGLTMEQKLISEEDLNEMKRLDHGSEVRSFSLGKPKVSEY 112
64 MXTENCV-----AKTKLANGTSMIVPKQKLSASYEKEKEL--CVKYFEQWSESD 112
113 TSTTGLVPCSATPTTFGLRAANGG-----QQRRTTSVQPTGLQEWLKMFSQWSGPE 167
113 QVEFVEHLISQCHYQHGHSYLYKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAA 172
168 KLLALDELIDSCPTQVKHMQVIEPQFQDFISLIP-----KELALYVLSFLEPKDLQA 223
173 ELVCKEWRVTSQDGLMWKLLIEMVVRTDSLRWGLAE-----RRGWQYLFKNKPPDGNAP 227
224 AQTCEYWLAEADNLLWRECKE-----EGIDEPLHIKRRK-----VIKPGFIHSP 269
228 PMSFVRALYPKIIODIETIESNWRGRHSQRHCRSETSGVYCLQYDDQKIVSGLRD 287
270 WKSAY-----IRQ--HRIDTNWRGELKSPKV-LKGHDDHVIITCLOFCGNNRIVSGSDN 320
288 TIKINDKNTLECKRIILTGHGTGSLVCLQYDERVITGSSDSTVRVMDVNTGEMLNTLIHHC 347
321 TLKWSAVTGKCLRTLVGHTGGVSSQMRDNIISGSDTRTLKVNNAETGECIHTLYGT 380

QY 288 TIKINDKNTLECKRIILTGHGTGSLVCLQYDERVITGSSDSTVRVMDVNTGEMLNTLIHHC 347
Db 321 TLKWSAVTGKCLRTLVGHTGGVSSQMRDNIISGSDTRTLKVNNAETGECIHTLYGT 380
QY 348 EAVLHLPNGMWTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSAS 407
Db 381 STVRCHMLHREKRVVSGSRDTRLVMDIETGQCL---HVLGMHVAARVCVQYDGRVYVSGA 437
QY 408 GDRTIKVNNTSTCEFVATLNGHKRGIACLOVRDLRVVSGSSDNTIRLWDIECCACLRVLE 467
Db 438 YDFWVKVMDPETETCLHTLOGHTNRVYSLOFDGHHVVGSLDTSIRWMDVETGNCIHTLT 497
QY 468 GHEELVRCIRPDNRKIVSGAYDGIKIVMDLVAALDPRAPAGTCLRLTV---BHSGRVFR 524
Db 498 GHQSLTSGMELKDNILVSGNADSTVKIWDIKTG-----QCLQTLQGNPKHQSAVTC 548
QY 525 LQDFEQIVSSSHDDTILWD 545
Db 549 LQFNKNFVITSSDDGTVKLWD 569
RESULT 10
US-09-328-877A-21
; Sequence 21, Application US/09328877A
; Patent No. US2002017187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328.877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6 myc tagged
; OTHER INFORMATION: homo sapien
US-09-328-877A-21
Query Match 21.2%; Score 644.5; DB 10; Length 626;
Best Local Similarity 30.1%; Pred. No. 1.8e-54;
Matches 169; Conservative 103; Mismatches 22; Indels 67; Gaps 15;
QY 8 LOEKALKFMSEREDCNGGPPPKIPEK--NSLRQ--TYNSCARLCLNQETVCLASTA 63
Db 53 MEQKLISEEDLNEMESLGLTMEQKLISEEDLNEMKRLDHGSEVRSFSLGKPKVSEY 112
QY 64 MXTENCV-----AKTKLANGTSMIVPKQKLSASYEKEKEL--CVKYFEQWSESD 112
Db 113 TSTTGLVPCSATPTTFGLRAANGG-----QQRRTTSVQPTGLQEWLKMFSQWSGPE 167
QY 113 QVEFVEHLISQCHYQHGHSYLYKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAA 172
Db 168 KLLALDELIDSCPTQVKHMQVIEPQFQDFISLIP-----KELALYVLSFLEPKDLQA 223
QY 173 ELVCKEWRVTSQDGLMWKLLIEMVVRTDSLRWGLAE-----RRGWQYLFKNKPPDGNAP 227
Db 224 AQTCEYWLAEADNLLWRECKE-----EGIDEPLHIKRRK-----VIKPGFIHSP 269
QY 228 PMSFVRALYPKIIODIETIESNWRGRHSQRHCRSETSGVYCLQYDDQKIVSGLRD 287
Db 270 WKSAY-----IRQ--HRIDTNWRGELKSPKV-LKGHDDHVIITCLOFCGNNRIVSGSDN 320
QY 288 TIKINDKNTLECKRIILTGHGTGSLVCLQYDERVITGSSDSTVRVMDVNTGEMLNTLIHHC 347
Db 321 TLKWSAVTGKCLRTLVGHTGGVSSQMRDNIISGSDTRTLKVNNAETGECIHTLYGT 380

348 EAVLHLRFNNGMWTCSKORSIAWDMASPTDITLRLVGHRAAVNVVDFDDKYIVSAS 407
381 STVRCMELHFKRVVSGSRDATLRWDIETGQCL---HVLGMHVAARVCQYDGRVVSQA 437
408 GDRITKWNSTCEFTVLNKHKGACIACQYRORLVVSGSDNTIRLWIECCACLRVLE 467
438 YDFMVKWDPEFETCLHTLOQHNRVVSLQFDGIHVVGSLDTSIRVWDVETGNCIHTLT 497
468 GHELVRCIRPDNKRIVSGYDGGKIKWDLVAALDPRAPAGTLCRLTLV---EHSGRVFR 524
498 GHQSLTSGMELKDNILVSGNADSTVKIWDIKTG-----QCLQTLOGPNKHQSANTC 548
525 LQDFEQIVSSSHDDTLIWD 545
549 LQFNKNFVITSSDDGTVKIMD 569

213-888-7
ance 7, Application US/09213888A
ic No. US20020164683A1
RAL INFORMATION:
ICANT: Gurney, Mark E.
ICANT: Li, Jinhe
ICANT: Pauley, Adele M.
ICANT: Pharmacia & Upjohn Company
E OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
E OF INVENTION: Encode them
REFERENCE: 6142
ENT APPLICATION NUMBER: US/09/213.888A
ENT FILING DATE: 1998-12-17
ER OF SEQ ID NOS: 27
WARE: Patent In Ver. 2.0
D NO 7
GTH: 540
E: PRT
ANISM: Homo sapiens
13-888-7

Match 21.1%; Score 640; DB 10; Length 540;
Local Similarity 30.8%; Pred. No. 4.1e-54;
es 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KFMSSREDNCNGEPPRKIIPEKNLSRQTYNSCARLCLNQETVCLASTAMKTENCVAKT 73
4 KLDHGEVRSFSLGKCKV-----SEYTSITGL-----VPCSA-----TPTTFGDL 45
74 KLANGTSMIVPKORKLSASVEKEKEL---CVKYFEQMSQSDQVEFVHLISQMHYQGH 131
46 RAANGQG-----QQRRTITSVQPTGLQEWLKFQMSGPEKLLALDELIDSCPTQVKH 100
132 INSVLKPMLODFITALPARGLDHIAENILSYLDAKSLCAELVCKEWEYVTS DGLMVK 191
101 MMQVIEPQDFDIFSLP-----KELALYVLSFLEPKDLQAQTCRYWRILAEDNLLWRE 156
192 LIERWVRTDSLWRGLAE-----RRGWQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETI 246
157 KCKE-----EGIDEPLHIKRRK-----VIKPGFIHSPWKSAY-----IRQ--HRI 194
247 ESNWRCGRHSIQRHCRSETSGYVCIQYDDOKITVSGLRDNTIKWKNLTLECKRLITGH 306
195 DTNWRREGELKSPKV-LKGHDHVTICLQFCGNRIVSGSDNTLTKVMSAVTGKCLRTLUGH 253
307 TGSVLCLOYDERVIITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMWTCSKD 366
254 TGGVWSSQWRNIIISGSDTDLTKVNAETGECIHTLYGHTSTVRCMHLHKEKRVVSGSRD 313
367 RSIADVMDASPTDITLRLVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTCBFVRTL 426
314 ATLRWDIETGQCL---HVLGMHVAARVCQYDGRVVSQYDGMVKVMDPETETCLHTL 370
427 NGHKEGACIACQYRDLVVGSSDNTIRLWIECCACLRVLEGEHELVRCIFDNKRIVSG 486

Db 371 QGHTNRVYSLOFDGIHVVGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 430
Qy 487 AYDGKIKWDLVAALDPRAPAGTLCRLTLV---EHSGRVFRLOFDEFQIYSSSHDDTILI 543
Db 431 NADSTVKIWDIKTG-----QCLQTLOGPNKHQSANTCLOFNKNFVITSSDDGTVKL 481
Qy 544 WD 545
Db 482 WD 483

RESULT 12
US-09-213-888-10
; Sequence 10, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213.888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-10

Query Match 21.1%; Score 640; DB 10; Length 540;
Best Local Similarity 30.8%; Pred. No. 4.1e-54;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

Qy 14 KFMSSREDNCNGEPPRKIIPEKNLSRQTYNSCARLCLNQETVCLASTAMKTENCVAKT 73
Db 4 KLDHGEVRSFSLGKCKV-----SEYTSITGL-----VPCSA-----TPTTFGDL 45
Qy 74 KLANGTSMIVPKORKLSASVEKEKEL---CVKYFEQMSQSDQVEFVHLISQMHYQGH 131
Db 46 RAANGQG-----QQRRTITSVQPTGLQEWLKFQMSGPEKLLALDELIDSCPTQVKH 100
Qy 132 INSVLKPMLODFITALPARGLDHIAENILSYLDAKSLCAELVCKEWEYVTS DGLMVK 191
Db 101 MMQVIEPQDFDIFSLP-----KELALYVLSFLEPKDLQAQTCRYWRILAEDNLLWRE 156
Qy 192 LIERWVRTDSLWRGLAE-----RRGWQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETI 246
Db 157 KCKE-----EGIDEPLHIKRRK-----VIKPGFIHSPWKSAY-----IRQ--HRI 194
Qy 247 ESNWRCGRHSIQRHCRSETSGYVCIQYDDOKITVSGLRDNTIKWKNLTLECKRLITGH 306
Db 195 DTNWRREGELKSPKV-LKGHDHVTICLQFCGNRIVSGSDNTLTKVMSAVTGKCLRTLUGH 253
Qy 307 TGSVLCLOYDERVIITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMWTCSKD 366
Db 254 TGGVWSSQWRNIIISGSDTDLTKVNAETGECIHTLYGHTSTVRCMHLHKEKRVVSGSRD 313
Qy 367 RSIADVMDASPTDITLRLVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTCBFVRTL 426
Db 314 ATLRWDIETGQCL---HVLGMHVAARVCQYDGRVVSQYDGMVKVMDPETETCLHTL 370
Qy 427 NGHKEGACIACQYRDLVVGSSDNTIRLWIECCACLRVLEGEHELVRCIFDNKRIVSG 486
Db 371 QGHTNRVYSLOFDGIHVVGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 430
Qy 487 AYDGKIKWDLVAALDPRAPAGTLCRLTLV---EHSGRVFRLOFDEFQIYSSSHDDTILI 543
Db 431 NADSTVKIWDIKTG-----QCLQTLOGPNKHQSANTCLOFNKNFVITSSDDGTVKL 481

```
544 WD 545
||
482 WD 483

13
328-877A-7
ence 7, Application US/09328877A
nt No. US20020177187A1
RAL INFORMATION:
LICANT: Gurney, Mark E.
LICANT: Li, Jinhe
LICANT: Pauley, Adele M.
LICANT: Pharmacia & Upjohn Company
E OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
E OF INVENTION: Encode Them
REFERENCE: 6142
CURRENT APPLICATION NUMBER: US/09/328, 877A
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
ID NO 7
ID NO 7
GTH: 540
E: PRT
ANISM: Homo sapiens
328-877A-7

/ Match 21.1%; Score 640; DB 10; Length 540;
Local Similarity 30.8%; Pred. No. 4.1e-54;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KFMNSREDNCNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTENCVAKT 73
4 KLDHGSSEVRSFSLGKKPKV-----SEYTTTGL-----VPCSA-----TPTTFGDL 45

74 KLANGTSSMIVPQKRLSASYEKEKEL--CVKYFEQWSESDDQVEFVEHLISOMCHYQGHG 131
46 RAANGQ-----QRRRITSVQPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVXH 100

132 INSVLKPMQLORDFTALPARGLDHIAENILSYLDAKSLCAAEVLCKEYRVTSDBGMLWK 191
101 MMQVIEPQFQDFISLLP-----KELALYVLSFLEPKDLLQAAQTCRYWRLAEDNLLWRE 156
192 LIERNVRTDSLWRGLAE-----RKGWGYLFKNKPPDGNAPPNSFYRALYPKIIQDIETI 246
157 KCKE-----EGIDEPLHIKRRK-----VIKPGFIHSPKMSAY-----IRQ--HRI 194
247 ESNWRCGRHSLQRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKINDKNTLECKRILTGH 306
195 DTNWRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGSDNTLVKMSAVTGKCLRTLUGH 253
307 TGSVLCLOYDERVLIITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366
254 TGGWSSQMRDNIISGSTDRTLKWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 313
367 RSIADVMDASPTDITLRRVLVGHRAAVNVVDFDKYIVSASGDRTIKVMNTSTCEFVRTL 426
314 ATLAVMDIETGQCL--HVLGMHVAARVCYQDGRVVSAYDFVMKVWMDPETETCLHTL 370
427 NGHKRGIAQLQYRDLVWSSDNTIRLWIDIECGACLRVLEGEHELVRCPFNKRIVSG 486
371 QGHTNRVYSLOFDGIHVWVGSLDTSIRVMDVETGNCIHTLTGHSLTSGMELKDNLIVSG 430
487 AYDGKIKVMDLVAALDPRAPAGTLCRLTLV--EHSGRVPRLOFDEFOIVSSSHDDTILI 543
431 NADSVTKIWDIKTG-----QCLQTLQPNKHQSNAVTCLOFNKNFVITSSDDGTVKL 481

544 WD 545
||
482 WD 483

RESULT 15
US-10-245-618-14
; Sequence 14, Application US/10245618
; Publication No. US20030143582A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
```

LICANT: Strohmaier, Heimo
LICANT: Spruck, Charles
LICANT: Sangfelt, Olie
UE OF INVENTION: HCD4 MODULATES CYCLIN B DEGRADATION
E REFERENCE: TSRI 779.2
RENT APPLICATION NUMBER: US/10/245,618
RENT FILING DATE: 2002-09-16
OR APPLICATION NUMBER: US 60/404,116
OR FILING DATE: 2002-08-15
OR APPLICATION NUMBER: US 60/322,947
OR FILING DATE: 2001-09-14
BER OF SEQ ID NOS: 48
WARE: FastSeq for Windows Version 4.0
ID NO 14
NGTH: 540
PE: PRT
GANISM: Artificial Sequence
ATURE:
HER INFORMATION: Synthesized
245-618-14
/ Match 21.1%; Score 640; DB 12; Length 540;
Local Similarity 30.8%; Pred. No. 4.1e-54;
ies 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
14 KPMSSSEREDCNGEPPEPKIIPKNSLRQTYNSCARLCLNQETVCLASTAMKTCNCVAKT 73
4 KLDHGSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPTFGDL 45
74 KLANGTSMIVPKQKLSASVEKEL--CVKYEQWSESDQVEFVEHLISQMCYQHG 131
46 RAANGCG-----QORRRITVSQPTGQEWLKNFQSWGPEKLLALDELIDSCPTGVKH 100
132 INSVLKPMLQRFITALPARGLDHIAENILSYLDAKSICAAELVCKEWRYVTSGLMVK 191
101 MNQVIEPQOFQDFISLLP-----KELALVLSFLBPKDLQAQTCRYWRILAEENLLWRE 156
192 LIERMVRTDLSLWRGLAE-----RRWGQYLFKNKFPDGNAPPNSFYALYPKIIOIETI 246
157 KCKE-----EGIDEPLHIKRR-----VIKPGFIHSPKSAV-----IRQ--HRI 194
247 ESNWRCGRHSIORIHCRSETSGVYCYQYDDQKIVSGLRNTIKIKDKNTLECKRLTGH 306
195 DTNWRGELKSPKV-LKGHDDHVIITCQFCGNRIVSGDDNTLRKWSAVTGKCLRTLUGH 253
307 TGSVLCQYDERVIITGSSDSTVRVWDYNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366
254 TGVWSSQWRDNIISGSTDRTLKVNAAETGECIHTLYGHTSTVRCMELHEKRVVSGRD 313
367 RSIATWDMASPTDITLRVLYCHRAAVNVDPDDKYIVSASGDRTIKWNNTSICEFVRTL 426
314 ATLRVWDIETGQCL---HVLGHWAAVRCVOYDGRVVRVSGAYDFMWKVWDPEPETCUHTL 370
427 NGHKRGIAQLQYRDLRVVSGSDNTIRLWDIECGACLRVLEGHBEELVRCIRFONKRVISG 486
371 QGHTNRVYSLQPDGHIHVVSGLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 430
487 AYDGKIKWDLVAALDPRAPACTLCRLTV--EHSGRVFLQDFEQIVSSSHDDTILI 543
431 NADSTVKIWDIKTG-----OCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTVKL 481
544 WD 545
482 WD 483

completed: October 22, 2003, 10:34:58
e : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

.ein - protein search, using sw model

October 22, 2003, 10:21:24 ; Search time 30 Seconds
(without alignments)
802.496 Million cell updates/sec

US-09-601-168b-2
score: 3034
e: 1 MDPAEVLQEKALKFNNSSP.....PAAQAPPPSPRTVTYISR 569

table: BLOSUM62
Gapop 10.0 , Gapext 0.5

id: 328717 seqs, 42310858 residues
number of hits satisfying chosen parameters: 328717

DB seq length: 0
DB seq length: 2000000000
rocessing: Minimum Match 0%
Listing first 45 summaries

- e : Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
 - 2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
 - 3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
 - 4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
 - 5: /cgn2_6/ptodata/2/iaa/PCtUS COMB.pcp.*
 - 6: /cgn2_6/ptodata/2/iaa/backfiles.pcp.*

red. No. is the number of results predicted by chance to have a
core greater than or equal to the score of the result being printed,
nd is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB ID	Description
2582.5	85.1	517	1	US-08-190-802A-30
2582.5	85.1	517	3	US-08-477-346-30
2582.5	85.1	517	4	US-08-473-089-30
2582.5	85.1	517	4	US-08-487-072A-30
545	18.0	640	4	US-08-177-165A-30
520	17.1	587	3	US-08-899-578-2
399	13.2	779	1	US-08-190-802A-32
399	13.2	779	3	US-08-477-346-32
399	13.2	779	4	US-08-473-089-32
399	13.2	779	4	US-08-487-072A-32
399	13.2	779	4	US-08-177-165A-29
392	12.9	732	4	US-08-914-999-8
354	11.7	409	2	US-08-283-917-3
354	11.7	409	2	US-08-961-716-3
354	11.7	410	2	US-08-283-917-9
354	11.7	410	2	US-08-961-716-9
339.5	11.2	409	1	US-08-190-802A-51
339.5	11.2	409	3	US-08-477-346-51
339.5	11.2	409	4	US-08-473-089-51
339.5	11.2	409	4	US-08-487-072A-51
337	11.1	1146	4	US-08-914-999-6
321.5	10.6	514	1	US-08-190-802A-66
321.5	10.6	514	3	US-08-477-346-66
321.5	10.6	514	4	US-08-473-089-66
321.5	10.6	514	4	US-08-487-072A-66
318	10.5	422	1	US-08-190-802A-52
318	10.5	422	3	US-08-477-346-52

Sequence 52, Appl
Sequence 52, Appl
Sequence 6, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 62, Appl
Sequence 3, Appl
Sequence 62, Appl
Sequence 62, Appl
Sequence 62, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 26, Appl
Sequence 2, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 18, Appl
Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-130-802A-30
; Sequence 30, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: MD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-190-802A-30

Query Match 85.1%; Score 2582.5; DB 1; Length 517;
Best Local Similarity 91.4%; Pred. No. 3.le-266;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;
18 SSERDCNNGEPFRKIPKNSLRQTYNSCARLCLNQETVCLASTMKTCNVAKTLAN 77

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-477-346-30

Query Match      85.1%; Score 2582.5; DB 3; Length 517;
Best Local Similarity 91.4%; Pred. No. 3.le-266; Indels 31; Gaps 2;
Matches 487; Conservative 7; Mismatches 8;

QY 18 SSEREDCNNGPPPKIIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAKTLAN 77
DB 13 ASEREDCNDRDEPPRKIIITEKNTRLQ-----TKLAN 42

QY 78 GTSSMIVPQKRLSASVEKEKELCVKYFEQWSESDDQVEFVEHLISOMCHYQHGHINSYLK 137
DB 43 GTSSMIVPQKRLSANVEKEKELCVKYFEQWSECDDQVEFVEHLISRCHYQHGHINTYLK 102

QY 138 PMLORDFITALPARGLDHAENILSYLDAKSLCAEVLCKEWRYVTSDGMLWKLIERMV 197
DB 103 PMLORDFITALPARGLDHAENILSYLDAKSLCSAEVLCKEWRYVTSDGMLWKLIERMV 162

QY 198 RTDSLWRGLAERRGWGOYLFKNKPPDGNAAPPNSFYRALYPKIIDTIESNWRRCGRHSL 257
DB 163 RTDSLWRGLAERRGWGOYLFKNKPPDGKTPPNSEFYRALYPKIIDTIESNWRRCGRHSL 222

QY 258 QRIHCSETSGKVCLQYDDOKIYVSGLRDNMTIKWDKNTLECKRILTGHTGSVLCLOYDE 317
DB 223 QRIHCSETSGKVCLQYDDOKIYVSGLRDNMTIKWDKNTLECKRILMGHTGSVLCLOYDE 282

QY 318 RVIIITGSSDSTVRVMDVNTGEMLNTLIHHCEAVLHLRFNNGMVTCSDRSIAVWDMASP 377
DB 283 RVIIITG-SDSTVRVMDVNTGEMLNTLIHHCEAVLHLRFNNGMVTCSDRSIAVWDMASA 341

QY 378 TDITLRVLVGHRRAAVNVDFDDKIYVSAGDRTIKVMNSTCFEVRTLNGHKRGIAQLQ 437
DB 342 TDITLRVLVGHRRAAVNVDFDDKIYVSAGDRTIKVMNSTCFEVRTLNGHKRGIAQLQ 401

QY 438 YRDRLVWGSSDNTIRLWDECGACLRVLEGHEELVRCIRFDNKRIYVSGAYDGKIKVWDL 497
DB 402 YRDRLVWGSSDNTIRLWDECGACLRVLEGHEELVRCIRFDNKRIYVSGAYDGKIKVWDL 461

QY 498 VAALDPRAPAGTICLRTLVEHSGRVFLQDFEQIVSSSHDDTTILIWFNLNDP 550
DB 462 VAALDPRAPAGTICLRTLVEHSGRVFLQDFEQIVSSSHDDTTILIWFNLNDP 514

RESULT 3
US-08-473-089-30
; Sequence 30 Application US/08473089
; Patent No. 6342368
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20008-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

n ce 30, Application US/08477346
t No. 6262023
PAL INFORMATION:
PLICANT: Mochly-Rosen, Daria
PLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
RESPONSE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
FOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
TORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
LECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 30:

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

RECENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,089

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

TORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.22

LECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

RMATION FOR SEQ ID NO: 30:

QUENCE CHARACTERISTICS:

LENGTH: 517 amino acids

TYPE: amino acid

TOPOLOGY: unknown

LECULE TYPE: peptide

POTHETICAL: NO

TI-SENSE: NO

IGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13

73-089-30

Match 85.1%; Score 2582.5; DB 4; Length 517;

Local Similarity 91.4%; Pred. No. 3.le-266;

es 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

18 SSEREDCNNGEPKPIPEKNSURQYNSCARLCLNQETVCLASTAMKTENCVAKTLAN 77

13 ASEREDCNDEPPRKIIITEKNTLRQ-----TKLAN 42

78 GTSSMIVPKORKLSASYEKEKELCVKYFEQWSESQDQVEFVHLISOMCHYQHGHSYLK 137

43 GTSSMIVPKORKLSANVEKEKELCVKYFEQWSECDQVEFVHLISRMCHYQGHINTYK 102

138 PMLQDFITLPPARGLDHIAENILSYLDAKSLCAELVCKEWYVTSQGLMKKLIERMV 197

103 PMLQDFITLPPARGLDHIAENILSYLDAKSLCAELVCKEWYVTSQGLMKKLIERMV 162

198 RTDSLWGLAERRGGWGYLFKNKPPDGNAPPNSFYRALYPKIIOETTESNWRGHRSL 257

163 RTDSLWGLAERRGGWGYLFKNKPPDGNAPPNSFYRALYPKIIOETTESNWRGHRSL 222

258 QRHCRSETSGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 317

223 QRHCRSETSGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 282

318 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMAS 377

283 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMAS 341

378 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRTIKVMNTSTCEFVRTLNHGKRGIAQLQ 437

342 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRTIKVMNTSTCEFVRTLNHGKRGIAQLQ 401

438 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHCELVLCRIRFNDKRIIVSGAYDGKIKVWDL 497

402 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHCELVLCRIRFNDKRIIVSGAYDGKIKVWDL 461

498 VAALDPRAPAGTICLRLTVHSGRVRLOPDEFOIVSSSHDDTILWDFLNDP 550

462 VAALDPRAPAGTICLRLTVHSGRVRLOPDEFOIVSSSHDDTILWDFLNDP 514

4

67-072A-30

nce 30, Application US/08487072A

t No. 6423684

RAL INFORMATION:

PLICANT: Mochly-Rosen, Daria

PLICANT: Ron, Dorit

TLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-487-072A-30

Query Match 85.1%; Score 2582.5; DB 4; Length 517;

Best Local Similarity 91.4%; Pred. No. 3.le-266;

Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

QY 18 SSEREDCNNGEPKPIPEKNSURQYNSCARLCLNQETVCLASTAMKTENCVAKTLAN 77

Db 13 ASEREDCNDEPPRKIIITEKNTLRQ-----TKLAN 42

QY 78 GTSSMIVPKORKLSASYEKEKELCVKYFEQWSESQDQVEFVHLISOMCHYQHGHSYLK 137

Db 43 GTSSMIVPKORKLSANVEKEKELCVKYFEQWSECDQVEFVHLISRMCHYQGHINTYK 102

QY 138 PMLQDFITLPPARGLDHIAENILSYLDAKSLCAELVCKEWYVTSQGLMKKLIERMV 197

Db 103 PMLQDFITLPPARGLDHIAENILSYLDAKSLCAELVCKEWYVTSQGLMKKLIERMV 162

QY 198 RTDSLWGLAERRGGWGYLFKNKPPDGNAPPNSFYRALYPKIIOETTESNWRGHRSL 257

Db 163 RTDSLWGLAERRGGWGYLFKNKPPDGNAPPNSFYRALYPKIIOETTESNWRGHRSL 222

QY 258 QRHCRSETSGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 317

Db 223 QRHCRSETSGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 282

QY 318 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMAS 377

Db 283 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMAS 341

QY 378 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRTIKVMNTSTCEFVRTLNHGKRGIAQLQ 437

Db 342 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRTIKVMNTSTCEFVRTLNHGKRGIAQLQ 401

QY 438 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHCELVLCRIRFNDKRIIVSGAYDGKIKVWDL 497

Db 402 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHCELVLCRIRFNDKRIIVSGAYDGKIKVWDL 461

```

498  VAALDPRAPAGTCLRTLVHSGRVFLQDFEQIVSSSHDDTILWFLNDP 550
|||||
462  VAALDPRAPAGTCLRTLVHSGRVFLQDFEQIVSSSHDDTILWFLNDP 514
|||||

7-165A-30
ce 30, Application US/09177165A
No. 6426205
L INFORMATION:
CANT: Tyers, Mike
CANT: Willems, Andrew
OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
OF INVENTION: DEPENDENT PROTEOLYSIS
REFERENCE: 11757.10USU1
NT APPLICATION NUMBER: US/09/177.165A
NT FILING DATE: 1998-10-22
APPLICATION NUMBER: 60/092,443
FILING DATE: 1998-07-10
APPLICATION NUMBER: 60/063,254
FILING DATE: 1997-10-24
R OF SEQ ID NOS: 50
ARE: Patentin Ver. 2.1
NO 30
TH: 640
: PRT
NISM: Saccharomyces cerevisiae
7-165A-30

Match 18.0%; Score 545; DB 4; Length 640;
Local Similarity 28.9%; Pred. No. 7.4e-49;
s 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

34  IPEKNSLRQYNSCARLCLN-QETVCLASTAMKTENCVAK-TKLANGTSSMIVPKQRKLS 91
|||
86  LPEYNFTKCYRHNPDIQFSPHTTACYQDKLKTQBEINANIAKLPLQEQSDIHIIISKYS 145
|||
92  ASYEKEKELCVKYPQWSESQVFEHLISQWCHVQHGHINSYKPLQDFITALPAR 151
|||
146  NSMDKIRKL-----ILGDIUSTCFPOLSYISLVTWHMIKIDFISLIP-- 188
|||
152  GLDHIAENILSYLDAKSLCAELVCKEYWRVTDGMLWKKLIERVWTDLSLWRGLABRR- 210
|||
189  ---QELSLKILSLDQCSLCNATRCRWKQLADDDRVVHMCQEH-----DEKC 236
|||
211  ---GWCQYLFKNK-----PDGNAPNSPYRALYFKIQTDIETIESNWRGCRHSLO 258
|||
237  PNGGWLPLHLHKRARIQNSTGSSNADIQTOTRPWKVYRFRPKVESNWRKG----- 291
|||
259  RIHCRSETSK---GVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRILTGHGTGSLCLO 314
|||
292  --HCRIQEPKGMHGVLTQLQFNRYLLFTGSDYSTIGIWDLFTGKILRLSGHSDGVKTL 349
|||
315  YDERVIITGSSDSTVRVDVNTGEMLNTLIHCEAVLHFRFNGMMVTCSDKRSIAWDM 374
|||
350  FDRKLKILGSLDKTIRVNYITGBCISTYRGHSDSVLSVDSYQKRVISGSADKTVKWHV 409
|||
375  ASPTDITLRLVLRVGHRAAVNVDFDDKYI---VSAAGDRITKYNWTSTCEFRVTLNGH--- 429
|||
410  ESRTCYYTLA---GHTEWVNCVKLHPKFSFCSCSDDTIRWWDITFNSCLKVFRGHVGQ 465
|||
430  KRGIACLQVRD---RLVWSSGS----- 448
|||
466  VQKIIFPTTKDVENIATDNTSDGSSPODDPTMTDGADESTPSNEQETVLDENIPYTHL 525
|||
449  -----DNTIRLWDIECGACLRVLEGHBEELVRCIRFONKRVISGAYDGKIKWDLVAALDP 503
|||
526  LSCGLDNTIKLMDVTKGCIQTQFCHVEGVWDAAFNRLISSGHGSKVWDLQSG--- 582
|||
504  RAPAGTCLRTLVHSGRVFLQDFEQIVSSSHDDTI 541
|||||
583  -----KCMHTF---NGR---RLQRETQHTQTSLGQKV 609
|||||

```

```

RESULT 6
US-08-899-578-2
; Sequence 2, Application US/08999578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JFW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-899-578-2

Query Match 17.1%; Score 520; DB 3; Length 587;
Best Local Similarity 28.8%; Pred. No. 3e-46; Indels 70; Gaps 17;
Matches 150; Conservative 78; Mismatches 222;

QY 72 KTKLANGTSSMI-----VPQQR--LSASYEKEKEL-----CVKYFEQWSESQVE 115
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
35 ESSYNSGSSSSYNADLSSSRPLQHKLDLSASPSRNNDLNPRVEHLIALFKDLSSAEQMD 94
QY 116 FVEHLISOMCHYQGHGHNISYLPKMLQRFITALPARGLDHFAENILSYLDAKSLCAAEVL 175
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 AFTEFLQESNMWNIQRLAIIPHPQRFPLSLPV-----ELGMKILNHLITGYDLLKVAQV 150
QY 176 CKEWRYVTSQMLWKKL--IERMV-----RTDSLWRGLAERRGQGYLFPNKPDPGNAP 227
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
151 SKXWKLISEIDKIKSLGLVEEFPKHPDPDTRVTGAWQGTAAAG-----VTIPDHIQIP 203
QY 228 PN-SFVEALYPKLIIDI-----ETIESNRCGRHSLQRIHCRSETSKGVYCL 273
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
204 CDLVNHRFLKQFGDIPERAAKDRLYLAUKIEKNWNANFMGSAV--LRGHEDHVTICM 262
QY 274 QYDDQKIVSGLRNTTIKWDKNTECKRLTGHGTGSLVLCLOYDE--RVITGSSDSTVRV 331
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
263 QIHDDVLVATGSDNTLKWVCIDKGEVMTYLVGHTGGVWTSQISQCGRIYIVSGSTDRTKV 322
QY 332 WDVNTGBMLNTLIHHCCEAVLHLRFNNGMWVTCSDRESIAVWDMASPTDITLRRVLVGHRA 391
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
323 WSTVDGSLLETLOCHTSTVRCVAMAGSLVLTGSRDITLVRWDVSEGRHLA---TLHGHH 379
QY 392 AVNVVDFDDKIYIVSAGDRTTIKWNTNSTCEFVRTLNGHKRGIAQLQYDR--LVVSGSD 449
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
380 AVRVCQPDGTTVYSGGYDFTKIKWAHATGCRITITLGHNNRVYSLLPESERSIVCSGLD 439

```

RESULT 8
 US-08-477-346-32
 Sequence 32, Application US/08477346
 Patent No. 6262023
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Moxtison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 779 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15
 US-08-477-346-32

Match 13.2%; Score 399; DB 3; Length 779;
Local Similarity 24.3%; Pred. No. 3.8e-33;
es 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTCNCVAKTKLANGTSSMIVPKQKLSASYEK-ELCVKVFQW 108
194 TTPLAKTTKINN-----NNNIADLIESKOSIISPEYLSDEIFSAINNPLHAYFK-- 244
109 SESQDVEFVHLISQCHYQHINSYLPKMLQORDPITALPARGLDHIAENILSYLDKAS 168
245 -----NLLFRLVANMDRSELSDLGLTKDKNLKRLDITSLPF-----EISLKFNYLQFED 294
169 LCAAEVLCKEWR-VTSDGMLWKKLI--ERMVTDLSLWGLAERRGQGYLFKNKPPDGN 225
295 IINSLGVSNQWNNKIIRKSTLSLWKKLIISNFV-----SPKGF 331
226 APPNSFYRALYPKIIQD-----IET--IESNWRGCRHSLQRIHCRSETSKGVYCLQYD 276
332 NSLNLKLSQKYPKLSQDRLRLSFLNIFILKNWYNPKFVPORTTLRGHMTSVITCLOFE 391
277 DQKIVGLRNTTIKWDKNTLECKRILTOHTGVSVCLOQYDE-RVITGSSDSTVRWVDN 335
392 DNVITGADDMIRVYDSINKKFLQLSGHDGGWALKYAHGGILVSGSDTRVWVDIK 451
336 TGEMLNTLIHCEAVLHRLFNNGMMVTCSDKRSIAVMDASPTDITLRVLVGHRAAVNV 395
452 KG-----CCT-----HVFEGHSTVRC 468
396 VDFDD---KYIVSASGDRITIKVWNT-----STCE-----FVRTLNG 428
469 LDIVEYKNIKYIVTGSRDNTLHWKLPKSSVPDHGEEHDYPLVFTHTPEENPFYGVLRG 528
429 HKRGIAQLQVRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN--KRIVSG 486
529 HMASVRTVSGHGNIVVSGSYDNTLIVMDVAQMKCLVILSGHTRDIYSTIYDHERKRCISA 588
487 AYDGKIKVMDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFRLQDFEQIVSS 535
589 SMDTTIRIWDLENINWNGECSYATNSASPCAKILGMYTLOQHTALVGLLRUSDKFLVSA 648
536 SHDDTILWDFLNDPAAQAEPPRSPRTYTY 566
649 AADGSIRGWD-AND-----YSRKFSY 668

Query Match 13.2%; Score 399; DB 4; Length 779;
Best Local Similarity 24.3%; Pred. No. 3.8e-33;
Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCLASTAMKTCNCVAKTKLANGTSSMIVPKQKLSASYEK-ELCVKVFQW 108
DB 194 TTPLAKTTKINN-----NNNIADLIESKOSIISPEYLSDEIFSAINNPLHAYFK-- 244
QY 109 SESQDVEFVHLISQCHYQHINSYLPKMLQORDPITALPARGLDHIAENILSYLDKAS 168
DB 245 -----NLLFRLVANMDRSELSDLGLTKDKNLKRLDITSLPF-----EISLKFNYLQFED 294
QY 169 LCAAEVLCKEWR-VTSDGMLWKKLI--ERMVTDLSLWGLAERRGQGYLFKNKPPDGN 225
DB 295 IINSLGVSNQWNNKIIRKSTLSLWKKLIISNFV-----SPKGF 331
QY 226 APPNSFYRALYPKIIQD-----IET--IESNWRGCRHSLQRIHCRSETSKGVYCLQYD 276
DB 332 NSLNLKLSQKYPKLSQDRLRLSFLNIFILKNWYNPKFVPORTTLRGHMTSVITCLOFE 391
QY 277 DQKIVGLRNTTIKWDKNTLECKRILTOHTGVSVCLOQYDE-RVITGSSDSTVRWVDN 335
DB 392 DNVITGADDMIRVYDSINKKFLQLSGHDGGWALKYAHGGILVSGSDTRVWVDIK 451
QY 336 TGEMLNTLIHCEAVLHRLFNNGMMVTCSDKRSIAVMDASPTDITLRVLVGHRAAVNV 395
DB 452 KG-----CCT-----HVFEGHSTVRC 468
QY 396 VDFDD---KYIVSASGDRITIKVWNT-----STCE-----FVRTLNG 428
DB 469 LDIVEYKNIKYIVTGSRDNTLHWKLPKSSVPDHGEEHDYPLVFTHTPEENPFYGVLRG 528
QY 429 HKRGIAQLQVRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN--KRIVSG 486
DB 529 HMASVRTVSGHGNIVVSGSYDNTLIVMDVAQMKCLVILSGHTRDIYSTIYDHERKRCISA 588
QY 487 AYDGKIKVMDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFRLQDFEQIVSS 535
DB 589 SMDTTIRIWDLENINWNGECSYATNSASPCAKILGMYTLOQHTALVGLLRUSDKFLVSA 648
QY 536 SHDDTILWDFLNDPAAQAEPPRSPRTYTY 566
DB 649 AADGSIRGWD-AND-----YSRKFSY 668

US-08-473-089-32

RESULT 10
US-08-487-072A-32
; Sequence 32, Application US/08487072A
; Patent No. 6423684
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron. Borit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265

RESPONSE ADDRESS:

ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA

ZIP: 20006-1812

MPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

TORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

LECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

RATION FOR SEQ ID NO: 32:

QUENCE CHARACTERISTICS:

LENGTH: 779 amino acids

TYPE: amino acid

TOPOLOGY: unknown

LECULE TYPE: protein

POTHEICAL: NO

TI-SENSE: NO

IGNAL SOURCE:

INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15

87-072A-32

Match 13.2%; Score 399; DB 4; Length 779;

Local Similarity 24.3%; Pred. No. 3.8e-33;

es 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTEKNCVAKTKLANGTSSMIVPKQKLSASYEK-ELCVKFFQW 108

194 TTPLAKTKTINN-----NNNIADIESKDSIIPEYLSDEIFSAINNPLHAYFK-- 244

109 SESQDQVEFVHLISQMHYGHINSYLKPEMLQORDITAPARGLDHIAENILSYLDAS 168

245 -----NLLPRLVANMDRSELSDLGLTIKDNKXDLTSLPF-----EISLKIFNYLQFED 294

169 LCAAEVLCKEYWR-VTSDGMLWKLI--ERMVTDLSLWGLAERRGQYLFKNKPPDGN 225

295 IINSLGVQSNWKIIKSTSLWKLLISENFV-----SPKGF 331

226 APPNSFYRALYPKIID-----IET--IESNRCGRHSLOHCRSETSGVYCLQYD 276

332 NSLNLKLSQKYPKLSQDRLRLSFLNIFILKWNYPKFPVQRTTLRGHMTSVITCLOFE 391

277 DQKIVSGLRNTIKWDKNTLECKRILTGTGVLCLQYDE-RVIITGSSDSTVRVMDVN 335

392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGGWALKYAHGGILVSGSTDRVYMDIK 451

336 TGEMLNTLIHCEAVLHRLFNNGMVTCSKDRSIAVMDNASPTDITLRLVLGHRAAVNV 395

452 KG-----CCT-----HVFEGHSTVRC 468

396 VFDD-----KYIVSAGDRTIKVWNT-----STCE-----FVRLTNG 428

469 LDIVEYKNIYIVTGSRDNTLHVWKLPKESVDPHGEEDYPLVPHTPPENPFVGVLRG 528

429 HKRGIAQLVDRDLVSGSDNTIRLWDIECGACLRLVEGHELVRCIRFDN--KRVISG 486

529 HMASVRTVSGHGNIVVSGSYDNTLIVMDVAQMKCLYLSGHTDRIVSTIYDHERKRCISA 588

487 AYDGKIKVWDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFLQDFQIVSS 535

Db 589 SMDTIRIWLENINWNGECSYATNSASPCAKILGAMTYLQHTALVGLRLSKFLVSA 648

Qy 536 SHDPTILWDLNDPAAQAEPSPSRPTY 566

Db 649 AADGSIRGWD-AND-----YSRKFSY 668

RESULT 11

US-09-177-165A-29

; Sequence 29, Application US/09177165A

; Patent No. 6426205

; GENERAL INFORMATION:

; APPLICANT: Tyers, Mike

; APPLICANT: Willems, Andrew

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN

; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS

; FILE REFERENCE: 11757.10USUI

; CURRENT APPLICATION NUMBER: US/09/177,165A

; CURRENT FILING DATE: 1998-10-22

; PRIOR APPLICATION NUMBER: 60/092,443

; PRIOR FILING DATE: 1998-07-10

; PRIOR APPLICATION NUMBER: 60/063,254

; PRIOR FILING DATE: 1997-10-24

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 29

; LENGTH: 779

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-177-165A-29

Query Match 13.2%; Score 399; DB 4; Length 779;

Best Local Similarity 24.3%; Pred. No. 3.8e-33;

Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

Qy 56 TVCLASTAMKTEKNCVAKTKLANGTSSMIVPKQKLSASYEK-ELCVKFFQW 108

Db 194 TTPLAKTKTINN-----NNNIADIESKDSIIPEYLSDEIFSAINNPLHAYFK-- 244

Qy 109 SESQDQVEFVHLISQMHYGHINSYLKPEMLQORDITAPARGLDHIAENILSYLDAS 168

Db 245 -----NLLPRLVANMDRSELSDLGLTIKDNKXDLTSLPF-----EISLKIFNYLQFED 294

Qy 169 LCAAEVLCKEYWR-VTSDGMLWKLI--ERMVTDLSLWGLAERRGQYLFKNKPPDGN 225

Db 295 IINSLGVQSNWKIIKSTSLWKLLISENFV-----SPKGF 331

Qy 226 APPNSFYRALYPKIID-----IET--IESNRCGRHSLOHCRSETSGVYCLQYD 276

Db 332 NSLNLKLSQKYPKLSQDRLRLSFLNIFILKWNYPKFPVQRTTLRGHMTSVITCLOFE 391

Qy 277 DQKIVSGLRNTIKWDKNTLECKRILTGTGVLCLQYDE-RVIITGSSDSTVRVMDVN 335

Db 392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGGWALKYAHGGILVSGSTDRVYMDIK 451

Qy 336 TGEMLNTLIHCEAVLHRLFNNGMVTCSKDRSIAVMDNASPTDITLRLVLGHRAAVNV 395

Db 452 KG-----CCT-----HVFEGHSTVRC 468

Qy 396 VFDD-----KYIVSAGDRTIKVWNT-----STCE-----FVRLTNG 428

Db 469 LDIVEYKNIYIVTGSRDNTLHVWKLPKESVDPHGEEDYPLVPHTPPENPFVGVLRG 528

Qy 429 HKRGIAQLVDRDLVSGSDNTIRLWDIECGACLRLVEGHELVRCIRFDN--KRVISG 486

Db 529 HMASVRTVSGHGNIVVSGSYDNTLIVMDVAQMKCLYLSGHTDRIVSTIYDHERKRCISA 588

Qy 487 AYDGKIKVWDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFLQDFQIVSS 535

Db 589 SMDTIRIWLENINWNGECSYATNSASPCAKILGAMTYLQHTALVGLRLSKFLVSA 648

Qy 536 SHDPTILWDLNDPAAQAEPSPSRPTY 566

[illegible]

14 961-716-3 Application US/08961716
Serial Information: at No. 5880272
Applicant: ADACHI, HIDEKI
Applicant: TSUJIMOTO, MASAFUMI
Applicant: INOUE, KEIJO
Applicant: ARAI, HIROYUKI
Title of Invention: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
Number of Sequences: 31
Correspondence Address:
Addressee: OBLON, SPIVAK, MCLELLAND, MAIER &
Addressee: NEUSTADT, P.C.
Street: 1755 S. Jefferson Davis Highway, Suite 400
City: Arlington
State: Virginia
Country: U.S.A.
Zip: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,716
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/283,917
FILING DATE: 03-AUG-1994
APPLICATION NUMBER: JP 209943/1993
FILING DATE: 03-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5880272man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2292-030-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Bos taurus
61-716-3

```

Query Match      11.7%; Score 354; DB 2; Length 409;
Best Local Similarity 29.2%; Pred. No.8e-29;
Matches 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;

QY 280 IVSGURDNITKIKWDKNTLECKRILTGHTGSVLCLQYDE--RVIIIGSSDSTVRVMDVNTG 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 MVASGEDATIKVWDYETGPFERTLKGTDSVDQISPDHSGKLLASCSDMTIKLWDFQGF 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 338 EMLNLTILHCEAVLHLP--NNGMMVTCSDRSIAWMDMASTDTILRRVLVGHRAAVNV 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 ECIRTMGHGDHNVSSVAINPNGHIVASRDXTKIMWEYOTGYCV--TFTGHREWVRM 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 396 V--DFDDKVIIVSASGRTTIKVNNTISCFVRTLNHGKRGIACLQYDR----- 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 VRPNQDGTLIASCNDQTVRVVWVATKECKAELREHEHVVEICSWAPESSYSSISEATGS 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 442 -----LVVSGSSDNTIILWDIECCACLRVLGHEELVRCIRFDN--KRIVSGAYD 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 EYTKSKGPGPFLLISGSRDXTIKWMDVSTQCMCLTVGHDNVVRGVLFHSGGKPFILSCADD 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 490 GKIKWDLVAALDPRAPAGTCLRLTILVHSGRVFLQFDEF--QIVSSSHDITLIWD 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 KTLIRVWDY-----KNKRCKWLTNAHEHFTSLDFHKTAPYVYVTVGSVDOTVKWVE 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Y Match 11.7%; Score 354; DB 2; Length 410;
Local Similarity 29.2%; Pred. No. 8.1e-29;
hes 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;
280 IVSGLRDNTIKIWKNTLECKXILTGHTGTVLCLOYDE--RVIIITGSSDSTVRVWDVNTG 337
123 MVSASEDAIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASCADMTIKLWDFQGF 162
338 EMLNTLIHCEAVLHLRF--NNGMMVTCSKDRSIATVMDMASPTDITLRRVLVGHRAAVNV 395
183 ECIRTMHGHNVSSVAIMPNGDHIVSASRDXTIKMWEVQTGYCV---KFTGHREWVRM 239
396 V--DFDDKYIIVSASGDRTIKVNWTSTCEFEVFTLNCHKRGIACTQYRDR----- 441
240 VRPNQDGTLIASCNDQTVRYVWVATKECAELREHEHVVECIWAPESSYSSISEATGS 299
442 -----LVVSGSSDNTIRLWDIEGACLRVLEGHEELVRCIRFDN--KRIVSGAYD 489
300 ETKKSGKPGPFLLSGSRDXTIKMWDVSTGCMCLMTLVGHDNWNVRGVLFHSGGKFILSCADD 359
490 GKIKVWDLVAALDPRAPAGTLCRLTVEHSGRVFRLOPDEF--QIVSSSHDDTILIWD 545
360 KTLRWVDY-----KNKRCMKTLANAHEHFTSLDFHKTAPYVVTGSDVDTVKVME 408

completed: October 22, 2003, 10:26:37
me : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

ein - protein search, using sw model

October 22, 2003, 10:14:07 ; Search time 104 Seconds
(without alignments)
1411.845 Million cell updates/sec

US-09-601-168b-2

score: 3034
e: 1 MDPAEAVLQEKALKEFMSSE.....PAQAEPFRSPRTYTYISR 569

table: BLOSUM62

Gapop 10.0 , Gapext 0.5

d: 830525 seqs, 258052604 residues

umber of hits satisfying chosen parameters: 830525

DB seq length: 0

DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

e: SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mmc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archheap.*

red. No. is the number of results predicted by chance to have a
core greater than or equal to the score of the result being printed,
nd is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB	ID	Description
2397	98.8	569	11	Q9QUI5	Q9qui5 mus musculus
2390	98.5	569	11	Q9Z159	Q9z159 mus musculus
2378	98.2	569	11	Q9R1G7	Q9r1g7 mus musculus
2395	78.9	529	11	Q8BY90	Q8by90 mus musculus
2380	78.4	563	11	Q8K022	Q8k022 mus musculus
2360	77.8	555	11	Q8CHD5	Q8chd5 mus musculus
2366	75.3	563	11	Q9Z3H0	Q9z3h0 mus musculus
2089.5	68.9	510	5	O44382	O44382 drosophila
2089.5	68.9	510	5	Q9VDE3	Q9vde3 drosophila
1498	49.4	430	5	Q9BJ54	Q9bj54 heterodera
1230	40.5	252	11	Q9Z2C7	Q9z2c7 mus musculus
661	21.8	1326	5	Q9VZF4	Q9vzf4 drosophila
640	21.1	553	4	Q9NUX6	Q9nux6 homo sapien
640	21.1	561	4	Q96R12	Q96r12 homo sapien
640	21.1	589	4	Q96LE0	Q96le0 homo sapien
640	21.1	627	4	Q96A16	Q96a16 homo sapien

Q969h0 homo sapien
Q9vhp4 mus musculu
Q9bv4 mus musculu
Q8uun3 xenopus lae
Q95zt0 caenothabdi
Q44083 caenothabdi
Q8ccs5 mus musculu
Q8xip2 podospora a
Q8xip4 podospora a
Q8xip3 podospora a
Q8xip5 podospora a
Q8z0r1 anabaena sp
Q8uun8 xenopus lae
Q9d4t2 mus musculu
Q8n136 homo sapien
Q8n776 homo sapien
Q8y109 anabaena sp
Q8z020 anabaena sp
Q96611 dictyosteli
Q8yz19 anabaena sp
Q8ysc0 anabaena sp
Q9c2l4 xenopus lae
Q9ptr5 gallus gall
Q8hxx0 macaca fasc
Q9gl51 sus scrofa
Q8avn1 xenopus lae
Q8kb12 chlorobium
Q96698 drosophila
Q810f4 dictyosteli

ALIGNMENTS

RESULT 1

Q9QUI5 PRELIMINARY; PRT; 569 AA.
AC Q9QUI5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)
DE (F-box-WD40 repeat protein 1).
GN BTBC OR FBXW1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9919275; PubMed=10097128;
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA Nakayama K.-I.;
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT ubiquitin ligase Skp1/Cul 1/F-box protein FWD1.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercuro F., Ben-Neriah Y.;
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase.";
RL Nature 396:590-594(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX Strausberg R.;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=21601157; PubMed=11735228;

ruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,
kayama K.-I.;
haxacterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of
enorhabdizans elegans SEL-10.";

enomics 78:214-222(2001)

BL; AF081887; AAL17755.1; -

BL; AF099932; AAD08701.1; -

BL; BC003989; AAH03989.1; -

BL; AF391190; AAL40929.1; -

BL; AF391178; AAL40929.1; JOINED.

BL; AF391179; AAL40929.1; JOINED.

BL; AF391180; AAL40929.1; JOINED.

BL; AF391181; AAL40929.1; JOINED.

BL; AF391182; AAL40929.1; JOINED.

BL; AF391183; AAL40929.1; JOINED.

BL; AF391184; AAL40929.1; JOINED.

BL; AF391185; AAL40929.1; JOINED.

BL; AF391186; AAL40929.1; JOINED.

BL; AF391187; AAL40929.1; JOINED.

BL; AF391188; AAL40929.1; JOINED.

BL; AF391189; AAL40929.1; JOINED.

D; MGI:1338871; Btrc.

terPro; IPR001810; F-box.

am; PF00646; F-box; 1.

am; PF00400; WD40; 7.

INTS; PR00320; GPROTEINBRPT.

ODOM; PD000018; WD40; 4.

IART; SM00320; WD40; 1.

OSITE; PS00181; FBOX; 1.

OSITE; PS00678; WD_REPEATS_1; 6.

OSITE; PS00082; WD_REPEATS_2; 7.

OSITE; PS00294; WD_REPEATS_REGION; 1.

gase; Repeat; WD repeat.

QUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Match 98.8%; Score 2997; DB 11; Length 569;

Local Similarity 98.6%; Pred. No. 9.6e-249;

es 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

1 MDPAAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLIA 60

1 MDPAAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLIT 60

61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYFQWSESQDQVEFVHL 120

61 STAMKTENCVAKAKLANGTSSMIVPKQKLSASYEKEKELCVKYFQWSESQDQVEFVHL 120

121 ISQCHYQHGHSYKPKMLQRFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 180

121 ISQCHYQHGHSYKPKMLQRFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 180

181 RVTSQDGLMVKKLIERWVRTDSLWRGLAERRGWQYLFKPKPPDGNAPPNSFYRALPKII 240

181 RVTSQDGLMVKKLIERWVRTDSLWRGLAERRGWQYLFKPKPPDGNAPPNSFYRALPKII 240

241 QDIETIESNRCGRHSIQRHCRSESKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

241 QDIETIESNRCGRHSIQRHCRSESKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

301 RILTGTGSLVCLQYDERVITGSSSDTVRVMDVNTGEMLTLIHCEAVLHLRFNNGMW 360

301 RILTGTGSLVCLQYDERVITGSSSDTVRVMDVNTGEMLTLIHCEAVLHLRFNNGMW 360

361 VTCSKDRSIADVMASTDTILRLVGHRAAVNVDFDKYIVSASGORTIKVNTSTC 420

361 VTCSKDRSIADVMASTDTILRLVGHRAAVNVDFDKYIVSASGORTIKVNTSTC 420

421 EFVRLTNGHKGACIQLQYRDLRVVSGSSDNTIRLWDIEGACLRVLEGEELVRCIRFDN 480

421 EFVRLTNGHKGACIQLQYRDLRVVSGSSDNTIRLWDIEGACLRVLEGEELVRCIRFDN 480

Qy 481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCITLVEHSGRVFRLQDFDFQIVSSHDDT 540

Db 481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCITLVEHSGRVFRLQDFDFQIVSSHDDT 540

Qy 541 ILIWDFLNDPAQAEPSPSRITYIISR 569

Db 541 ILIWDFLNDPAQAEPSPSRITYIISR 569

RESULT 2

Q9Z159

ID Q9Z159

AC Q9Z159

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Beta-transducin repeat containing protein.

GN BTRC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

FN [1]

RP SEQUENCE FROM N.A.

EX MEDLINE=99145465; PubMed=9990853;

RA Spencer E., Jiang J., Chen Z.J.;

RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein

RT Slimb/beta-TrCP.";

RL Genes Dev. 13:284-294(1999).

DR EMBL; AF112979; AAD04181.1; -

DR MGD; MGI:1338871; Btrc.

DR InterPro; IPR001810; F-box.

DR InterPro; IPR001680; WD40.

DR Pfam; PF00646; F-box; 1.

DR PRINTS; PR00320; GPROTEINBRPT.

DR ProDom; PD000018; WD40; 4.

DR SMART; SM00356; FBOX; 1.

DR SMART; SM00320; WD40; 7.

DR PROSITE; PS00181; FBOX; 1.

DR PROSITE; PS00678; WD_REPEATS_1; 6.

DR PROSITE; PS00082; WD_REPEATS_2; 7.

DR PROSITE; PS00294; WD_REPEATS_REGION; 1.

KW Repeat; WD repeat.

SQ SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;

Query Match 98.5%; Score 2990; DB 11; Length 569;

Best Local Similarity 98.4%; Pred. No. 3.8e-248;

Matches 560; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MDPAAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLIA 60

Db 1 MDPAAVLQKALKFNMSSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLIT 60

Qy 61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYFQWSESQDQVEFVHL 120

Db 61 STAMKTENCVAKAKLANGTSSMIVPKQKLSASYEKEKELCVKYFQWSESQDQVEFVHL 120

Qy 121 ISQCHYQHGHSYKPKMLQRFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 180

Db 121 ISQCHYQHGHSYKPKMLQRFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEY 180

Qy 181 RVTSQDGLMVKKLIERWVRTDSLWRGLAERRGWQYLFKPKPPDGNAPPNSFYRALPKII 240

Db 181 RVTSQDGLMVKKLIERWVRTDSLWRGLAERRGWQYLFKPKPPDGNAPPNSFYRALPKII 240

Qy 241 QDIETIESNRCGRHSIQRHCRSESKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

Db 241 QDIETIESNRCGRHSIQRHCRSESKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

Qy 301 RILTGTGSLVCLQYDERVITGSSSDTVRVMDVNTGEMLTLIHCEAVLHLRFNNGMW 360

Db 301 RILTGTGSLVCLQYDERVITGSSSDTVRVMDVNTGEMLTLIHCEAVLHLRFNNGMW 360

```

361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVYSASGDRITIKVNTSTC 420
|||||
361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVYSASGDRITIKVNTSTC 420
|||||
421 EFVRLTNGHKGRIACIQYDRDLVVGSSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
421 EFVRLTNGHKGRIACIQYDRDLVVGSSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||
541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||

```

3

R1G7; PRELIMINARY; PRT; 569 AA.

R1G7; PRELIMINARY; PRT; 569 AA.

-MAY-2000 (TrEMBLrel. 13, Created)

-MAY-2000 (TrEMBLrel. 13, Last sequence update)

-OCT-2002 (TrEMBLrel. 22, Last annotation update)

ta-transducin repeat-containing protein.

*C.

us musculus (Mouse).

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

}

} QUENCE FROM N.A.

nston J., Ellledge S.J., Harper J.W.;

bmitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

BL; AF10396; AA041025.1; -

D; MG1:1338871; Btrc.

terPro; IPR001810; F-box.

am; PF00646; F-box; 1.

am; PF00400; WD40; 7.

INTS; PR00320; GPROTEINRPT.

cdcm; PD000018; WD40; 4.

ART; SM00256; FBOX; 1.

ART; SM00320; WD40; 7.

OSITE; PS50181; FBOX; 1.

OSITE; PS00678; WD_REPEATS_1; 6.

OSITE; PS00082; WD_REPEATS_2; 7.

OSITE; PS00294; WD_REPEATS_REGION; 1.

peat; WD repeat.

QUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

Match 98.2%; Score 2978; DB 11; Length 569;

Local Similarity 97.9%; Pred. No. 4.1e-247;

es 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

1 MDPABAVLQEKALPFMNSRERDCNNGEPPRKIIPEKNSLRQYNSCARLCLNQETVCLA 60

1 MDPABAVLQEKALPFMNSRERDCNNGEPPRKIIPEKNSLRQYNSCARLCLNQETVCLT 60

61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKELCVKYFEQWSESDDQVEFVEHL 120

61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKELCVKYFEQWSESDDQVEFVEHL 120

121 ISQCHYQHGHSINLYKPLMQORDITAPARGLDHIAENILSYLDAKSLCAAELVCKEY 180

121 ISQCHYQHGHSINLYKPLMQORDITAPARGLDHIAENILSYLDAKSLCAAELVCKEY 180

181 RVTSQGLMWKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240

181 RVTSQGLMWKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240

241 QDIETIESNRCGRHSRQRIHCRSETSGKYVCLQYDDQKIVSGRLDNTIKIWDKNTLECK 300

```

Db 241 QDIETIESNRCGRHSRQRIHCRSETSGKYVCLQYDDQKIVSGRLDNTIKIWDKNTLECK 300
|||||
Qy 301 RILTGTGTVLCLOYDERVITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
|||||
Db 301 RILTGTGTVLCLOYDERVITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
|||||
Qy 361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVYSASGDRITIKVNTSTC 420
|||||
Db 361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVYSASGDRITIKVNTSTC 420
|||||
Qy 421 EFVRLTNGHKGRIACIQYDRDLVVGSSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
Db 421 EFVRLTNGHKGRIACIQYDRDLVVGSSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
Qy 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
Db 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
Qy 541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||
Db 541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||

```

RESULT 4

```

ID Q8BY90 PRELIMINARY; PRT; 529 AA.
AC Q8BY90;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE F-box/WD-repeat protein 1B.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002)."
DR EMBL: AK041532; BAC30975.1; -
SQ SEQUENCE 529 AA; 60812 MW; D2F382457FD90080 CRC64;

```

Query Match 78.9%; Score 2395; DB 11; Length 529;
 Best Local Similarity 78.9%; Pred. No. 4.6e-197;
 Matches 449; Conservative 43; Mismatches 37; Indels 40; Gaps 3;

```

Qy 1 MDPABAVLQEKALPFMNSRERDCNNGEPPRKIIPEKNSLRQYNSCARLCLNQETVCLA 60
|||||
Db 1 MEP-DSVIEDTKTIELMTSWMEDQNEDESPKK----- 31
|||||

```

```

Qy 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKELCVKYFEQWSESDDQVEFVEHL 120
|||||
Db 32 -----SALWQISNGTSSVIVSRKRPSGNYQKEKOLCIKYFDQWSSDDQVEFVEHL 82
|||||

```

```

Qy 121 ISQCHYQHGHSINLYKPLMQORDITAPARGLDHIAENILSYLDAKSLCAAELVCKEY 180
|||||
Db 83 ISRCHYQHGHSINLYKPLMQORDITAPARGLDHIAENILSYLDAKSLCAAELVCKEY 142
|||||

```

```

Qy 181 RVTSQGLMWKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240
|||||
Db 143 RVISEGMLWKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 200
|||||

```

```

Qy 241 QDIETIESNRCGRHSRQRIHCRSETSGKYVCLQYDDQKIVSGRLDNTIKIWDKNTLECK 300
|||||
Db 201 QDIETIESNRCGRHSRQRIHCRSETSGKYVCLQYDDQKIVSGRLDNTIKIWDKNTLECK 260
|||||

```

```

Qy 301 RILTGTGTVLCLOYDERVITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
|||||

```

```
261 KVLGTGTVLCQYDERVITVSGSDTVRVWDVNTGEVNTLIHHNEAVLHRTSNGLM 320
262 VTCSKORSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKYIVSASGDRITKAVNTSTC 420
263 VTCSKORSIAVWDMASATDITLRVLVGHRAAVNVVDFDDKYIVSASGDRITKAVNTSTC 380
264 EFVRLTLNGHKGACIQYDRRLVVGSSSDNTIRLWDIECGACLRVLVEGHEELVRCIRFDN 480
265 EFVRLTLNGHKGACIQYDRRLVVGSSSDNTIRLWDIECGACLRVLVEGHEELVRCIRFDN 440
266 KRIVSGAYDGIKWMDLVAALDPRAPAGTLCIRTLVEHSGRVFRLOPDEFOIVSSSHDT 540
267 KRIVSGAYDGIKWMDLVAALDPRAPAGTLCIRTLVEHSGRVFRLOPDEFOIVSSSHDT 500
268 ILIWDFLNDPAQAEPSPSRVTYISR 569
269 ILIWDFLNVPPSAQNETRSPSRVTYISR 529
270 K022: PRELIMINARY; PRT; 563 AA.
271 -OCT-2002 (TrEMBLrel. 22, Created)
272 -OCT-2002 (TrEMBLrel. 22, Last sequence update)
273 -MAR-2003 (TrEMBLrel. 23, Last annotation update)
274 box and WD-40 domain protein 1B.
275 s musculus (Mouse).
276 karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
277 mmalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
278 BL_TaxID=10090;
279 QUENCE FROM N.A.
280 SSUE=Kidney;
281 rausberg R.;
282 mitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
283 BL; BC034261; AA3434261.1; -.
284 terPro; IPR001810; F-box.
285 terPro; IPR001680; WD40.
286 am; PF00646; F-box; 1.
287 am; PF00400; WD40; 7.
288 INTS; PR00320; GPROTEINRPT.
289 ODom; PD000018; WD40; 4.
290 IART; SM00256; FBOX; 1.
291 IART; SM00320; WD40; 7.
292 OSITE; PS50181; FBOX; 1.
293 OSITE; PS00678; WD REPEATS_1; 5.
294 OSITE; PS50082; WD REPEATS_2; 7.
295 OSITE; PS50294; WD_REPEATS_REGION; 1.
296 eat; WD repeat.
297 QUENCE 563 AA; 64481 MW; 74D07E6BB850ABED CRC64;
298 Match 78.4%; Score 2380; DB 11; Length 563;
299 Local Similarity 78.1%; Pred. No. 9.7e-196;
300 es 454; Conservative 46; Mismatches 51; Indels 30; Gaps 7;
301 1 MDAEVLQEKALAFMNSERE---DCVNGEPPRKIPKNSLRQTVNSCARLQNG--E 55
302 1 MEP-DSVIEDTKIELMCSVPSRLWLGAN-----LVESMCAIASCLOSMP 44
303 56 TVCLASTA-MKTEN-----CVAKTLANGTSSMIVPKRKLSASYEKEKELCVKYPEQW 108
304 45 VRCLQNTSMVEDQNEDESPKKSALWQISNGTSSVIVSRKRPSEGNVQKEDLCIKYFDQW 104
305 109 SESQDQVEVHLISQMCHYQGHGINSYKLPMLQDFITALPARGLDHIAENILSYLDAKS 168
306 105 SESQDQVEVHLISRMCHYQGHGINSYKLPMLQDFITALPEQGLDCHIAENILSYLDAKS 164
307 169 LCAELVCKEWYRVTSDGMLWKILIERWRTDSLWRGLAERRGQVLFKNKPPDGNAPP 228
308 165 LCAELVCKEWQVISEGMLWKILIERWRTDPLWGLSERGWDQVLFKNRPTDG--PP 222
```

```
229 NSFYRALPKIIQDIETIESNWRCCGRHSQRHCRSETSGVYCLQYDDQKIVSGLRDMT 288
223 NSFYSLPKIIQDIETIESNWRCCGRHNLQICRSNSKGVYCLQYDDDKIISGLRDS 282
289 IKWDKNTLECKILLTGHTGSLVLCQYDERVITVSGSDTVRVWDVNTGEMTLNLIHCE 348
283 IKWDKSSLECLKVLGTGHTGSLVLCQYDERVITVSGSDTVRVWDVNTGEMTLNLIHNE 342
349 AVLHLRFNNGMVTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKYIVSASG 408
343 AVLHLRFNGLMVTCSKDRSIAVWDMASATDITLRVLVGHRAAVNVVDFDDKYIVSASG 402
409 DRIKWNTSTCEPVTILNGHKGACIQYDRRLVVGSSSDNTIRLWDIECGACLRVLEG 468
403 DRIKWNTSTCEPVTILNGHKGACIQYDRRLVVGSSSDNTIRLWDIECGACLRVLEG 462
469 HEELVRCIRFDNKRIVSGAYDGIKWMDLVAALDPRAPAGTLCIRTLVEHSGRVFRLOP 528
463 HEELVRCIRFDNKRIVSGAYDGIKWMDLVAALDPRAPAGTLCIRTLVEHSGRVFRLOP 522
529 EFQIVSSSHDDTILWDFLNDPAQAEPSPSRVTYISR 569
523 EFQIISSHDDTILWDFLNVPPSAQNETRSPSRVTYISR 563
309 RESULT 6
310 ID Q8CHDS PRELIMINARY; PRT; 555 AA.
311 AC Q8CHDS;
312 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
313 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
314 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
315 DE MKIAA0696 protein (Fragment).
316 GN MKIAA0696.
317 OS Mus musculus (Mouse).
318 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
319 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
320 OX NCBI_TaxID=10090;
321 RN [1]
322 RP SEQUENCE FROM N.A.
323 RC TISSUE=Brain;
324 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Hara Y., Nagase T.,
325 Ohara O., Koga H.;
326 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
327 I. The complete nucleotide sequences of 100 mouse KIAA-homologous
328 cDNAs identified by screening of terminal sequences of cDNA clones
329 randomly sampled from size-fractionated libraries."
330 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
331 DR EMBL; AB093260; BAC41444.1; -.
332 FT NON TER 1
333 SQ SEQUENCE 555 AA; 63287 MW; 6025496DBAC6A489 CRC64;
334 Query Match 77.8%; Score 2360; DB 11; Length 555;
335 Best Local Similarity 81.9%; Pred. No. 5e-194;
336 Matches 443; Conservative 38; Mismatches 46; Indels 14; Gaps 3;
337 29 PPRKIIPEKNSLRQTVNSCARLQNGTVCLASTAMKTCNCVAKTLANGTSSMIVPKOR 88
338 29 PPGFIAAE---RRERGGAAGAMEPDSV-----IEDKTIELMISNGTSSVIVSRKR 76
339 89 KLSASYEKEKELCVKYPEQWSESDQVEFVHLISQMCHYQGHGINSYKLPMLQDFITAL 148
340 77 PSEGNVQKEDLCIKYFDQWSESDQVEFVHLISRMCHYQGHGINSYKLPMLQDFITAL 136
341 149 PARGLDHIAENILSYLDAKSICAEELVCKEWYRVTSDGMLWKILIERWRTDSLWRGLAE 208
342 137 PEQGLDCHIAENILSYLDAKSICAEELVCKEWQVISEGMLWKILIERWRTDPLWGLSE 196
343 209 RRGWGQVLFKNKPPDGNAPPNSFYRALPKIIQDIETIESNWRCCGRHSQRHCRSETSK 268
344 197 RRGWDQVLFKNRPTDG--PNSFYSLPKIIQDIETIESNWRCCGRHNLQICRSNSK 254
345 269 GYVCLQYDDQKIVSGLRDMTIIKWMDKNTLECKILLTGHTGSLVLCQYDERVITVSGSDST 328
```

```

255 GVTCLQYDDDKIISGLRDSIKIWKSSLECLVLTGHTGSLVCLQYDERVIVTSSDST 314
329 VRVMDVNTGBMLNTLIHHCBAVLHLRPNNGMMVTCSDRSIAVWDMASPTDITLRLVLVG 388
315 VRVMDVNTGBVLTNLIHNEAVLHLAFSNGLMVTCSDRSIAVWDMASATDITLRLVLVG 374
389 HRAAVNVDPDDKVIIVSASGDRITIKYWNSTCEVFTLNHKGKGIACLOYRDLVVSGSS 448
375 HRAAVNVDPDDKVIIVSASGDRITIKYWNSTCEVFTLNHKGKGIACLOYRDLVVSGSS 434
449 DNTIRLWDIECGACLRVLEGHEBELVRCIRFNDKRIIVSGAYDGKIKWMDLVAALDPRAPAG 508
435 DNTIRLWDIECGACLRVLEGHEBELVRCIRFNDKRIIVSGAYDGKIKWMDLVAALDPRAPAS 494
509 TLLCLRTLVHSGRVFRLOQDFEFOIVSSSHDDTILWDFNDPAAQAEPPSPRTTYIS 568
495 TLLCLRTLVHSGRVFRLOQDFEFOIVSSSHDDTILWDFNVPPSAQNSTRSPRTTYIS 554
569 R 569
555 R 555

23HO PRELIMINARY; PRT; 563 AA.
23HO:
-DEC-2001 (TrEMBLrel. 19, Created)
-DEC-2001 (TrEMBLrel. 19, Last sequence update)
-MAR-2003 (TrEMBLrel. 23, Last annotation update)
box/wd40 repeat-containing protein HOS.
XXWL.
is musculus (Mouse)
Karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
} _TaxID=10090;
RAIN=FVB/N;
atia N., Herter J.R., Siaga T.J., Fuchs S.Y., Spiegelman V.S.;
ouse homolog of HOS (mHOS) is overexpressed in skin tumors and
licated in constitutive activation of NF-kappaB;
mitted (JUN-2001) to the ENBL/GenBank/DBJ databases.
BL; AY038079; AAK2095.1; -.
D; MGI:2144023; Fbxwlb.
terPro; IPR001810; F-box.
am; PF00646; F-box; 1.
am; PF00400; WD40; 7.
INTS; PR00320; GPROTEINBRPT.
ODom; PD000018; WD40; 4.
IART; SM00256; FBOX; 1.
IART; SM00320; WD40; 7.
OSITE; PS50181; FBOX; 1.
OSITE; PS00678; WD REPEATS_1; 5.
OSITE; PS50082; WD REPEATS_2; 7.
OSITE; PS50294; WD REPEATS_REGION; 1.
beat; WD repeat.
QUENCE 563 AA; 64741 MW; 9AB562F3FF5E3496 CRC64;

Match 75.3%; Score 2286; DB 11; Length 563;
Local Similarity 75.7%; Pred. No. 1.2e-187;
es 440; Conservative 46; Mismatches 65; Indels 30; Gaps 7;

1 MDPBAVLQEKALFPMNSERE---DQNGPPPKIIPKNSLRQTNSCARCLNQ--E 55
1 MEP-DSVLTETIEMCSVPASLMUGCN-----LVSMALSCIQMPS 44

56 TVCLASTA-MKTEN-----CVAKTKLANGTSSMIVPKRKLSAYEKEKELCVKYPQW 108
45 VRCLQNTSMWEDNEDESPKKSALWQISNGTSSVIVSEKRPSEGNQKENDLCIKYFDQW 104

109 SEDQVEFVEHLISOMCHYOHGHINSYKPMLOQDFITALPARGLDHIAENILSYLDAKS 168
105 SEDQVEFVEHLISRMCHYOHGHINSYKPMLOQDFITALPEGLDHAENILSYLDARS 164
169 LCAAEVCKEWMYRTSDGMLWKKLIERMVTRTDSLWGLAERRGWGYLFKNKPPDGNAPP 228
165 LCAAEVCKEWMQVISEGMLWKKLIERMVTRTDLWKLSERRGWQYLFKNRPTDG--PP 222
229 NSFVRALYPKIIODIETIESNMECGRHSIORHCRSETSGVYVCLQYDDOKIVSGLRDNT 288
223 NSFVRSYLPKIIODIETIESNMECGRHNLIQICRSSENSKGVYVCLQYDDOKIISGLRDS 282
289 IKIWDKNTLBECKRILTGHGTGSLVCLQYDERVIVTSSDSTVRVWDMVNTGEMNTLIHCE 348
283 IKIWDKSSLECLVLTGHTGSLVCLQYDERVIVTSSDSTVRVWDMVNTGEVNTLIHNE 342
349 AVLHLRFNNGMMVTCSDRSIAVWDMASPTDITLRLVGHRAAVNVDPDDKVIIVSASG 408
343 AVLHLRFNGLMVTCSKRSIAVWDMAFCHRYHFTPCSGWPRAAVNVDPDDKVIIVSASG 402
409 DRTIKVWNTSTCEVFTLNHKGKGIACLOYRDLVVSGSSDNTIRLWDIECGACLRVLEG 468
403 DRTIKWSTSTCEVFTLNHKGKGIACLOYRDLVVSGSSDNTIRLWDIECGACLRVLEG 462
469 HEELVRCIRFNDKRIIVSGAYDGKIKWMDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQD 528
463 HEELVRCIRFNDKRIIVSGAYDGKIKWMDLVAALDPRAPASTLCRLTLVHSGRVFRLOQD 522
529 EFOIVSSSHDDTILWDFNDPAAQAEPPSPRTTYISR 569
523 EFOIVSSSHDDTILWDFNVPPSAQNSTRSPRTTYISR 563

RESULT 8
O44382 PRELIMINARY; PRT; 510 AA.
ID C04382;
AC C04382;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE SLIMB.
GN SLMB OR SLIMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98121115; PubMed=9461217;
RA Jiang J., Struhl G.;
RT "Regulation of the Hedgehog and Wntless signalling pathways by the F-
RT box/wd40-repeat protein Slmb.";
RL Nature 391:493-496(1998).
DR EMBL, AF032878; AAC38852.1; -.
DR FlyBase; FBgn0023423; slmb.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 4.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD REPEATS_1; 5.
DR PROSITE; PS50082; WD REPEATS_2; 7.
DR PROSITE; PS50294; WD REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; DB0243D3730A5E8 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;
Best Local Similarity 78.3%; Pred. No. 7.7e-171;
Matches 394; Conservative 48; Mismatches 54; Indels 7; Gaps 4;

```

```

64 MKTENCVAKTCLANG---TSSMIVPKORKLSAS--YEKEKELCVKYFEOWSESQDQVEFVE 118
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Ovary, and Imaginal disks;
RX MEDLINE=20245299; PubMed=10781936;
RA Miletich I., Limbourg-Bouchon B.;
RT "Drosophila null limb clones transiently deregulate Hedgehog-
independent transcription of wingless in all limb discs, and induce
decapentaplegic transcription linked to imaginal disc regeneration.";
RL Mech. Dev. 93:15-26(2000).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003733; AAF55853.1; -.
DR EMBL; AF222924; AAF63214.1; -.
DR EMBL; AF222923; AAF63213.1; -.
DR EMBL; AY118898; AAF50758.1; -.
DR FlyBase; FB00023423; slmb.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PRODOM; PD000018; WD40; 4.
DR PROSITE; PS0181; FBOX; 1.
DR PROSITE; PS0678; WD_REPEATS_1; 5.
DR PROSITE; PS0082; WD_REPEATS_2; 7.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; F4D5DF126F58A012 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;
Best Local Similarity 78.3%; Pred. No. 7.7e-171;
Matches 394; Conservative 49; Mismatches 53; Indels 7; Gaps 4;

QY 64 MKTENCVAKTCLANG---TSSMIVPKORKLSAS--YEKEKELCVKYFEOWSESQDQVEFVE 118
Db 4 METDKINDETN-SNAQAFTTTMLYDPVRKXDSPTTYOTERELCFQYFTQWSESQDQVE 62
QY 119 HLISOMCHYQHGHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKE 178
Db 63 HLLSRMCHYQHGHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKE 122
QY 179 WYRTSDGMWKKLIERWRTDSLWRGLAERRGWGYLFKNKPPDGNAPNSFVRLYPK 238
Db 123 WLRVISEGMLWKLIERKVRTDSLWRGLAERRNMQYLFRPQG-QTORPHSFHRELFPK 181

```


1 239 IIQDIETIESNWRGRRHSIORIHCRRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLE 298
182 IMNDIDSIENNRWTRGHMLLRINCRRSENKGVYCLQYDDQKIVSGLRDNTIKIWDRTDQ 241
299 CKRILTCTGTSVLCLOYDERRVITGSSDSTVRVWDVNTGEMLNTLIHHCVAVLHRRKNG 358
242 CVKTLMGHTGSLVLCLOYDQKVIISGSSDSTVRVWDVNTGEMVNTLIHHCVAVLHRRFNG 301
359 MMVTCSDRSIAVMDMASPTDILRRVLVGHRAAVVVDPDDKVIIVSASGDRTIKVWNTS 418
302 MMVTCSDRSIAVMDWNTSPSEITLRRVLVGHRAAVVVDDEKVIIVSASGDRTIKVWSTS 361
419 TCEVRTLNGHKGIGIACILOYRRLVVGSSDNTLRLWDIECGACLRVLGHEELVRCIRF 478
362 SCEVRTLNGHKGIGIACILOYRRLVVGSSDNTLRLWDIECGACLRVLGHEELVRCIRF 421
479 DNKRIVSGAYDGKIKWMDLVAALDPRAPAGTLCRLTIVHSGRVRFLQDFEQIVSSSHD 538
422 DTKRIVSGAYDGKIKWMDLVAALDPRASNTLCNTLVAHTGRVRLQDFEQIVSSSHD 481
539 DTLIWDFLNDPAAQAEPPRSPS 561
482 DTLIWDFLNTPNENKTKGTFS 504
BU54 PRELIMINARY; PRT; 430 AA.
JUN-2001 (TrEMBLrel. 17, Created)
-JUN-2001 (TrEMBLrel. 17, Last sequence update)
-OCT-2002 (TrEMBLrel. 22, Last annotation update)
a-transducin repeat-containing protein (Fragment).
eroderata glycines (Soybean cyst nematode).
aryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
lenchoidea; Heteroderidae; Heteroderinae; Heterodera.
BI_TaxID=51029;
SEQUENCE FROM N.A.
alava E.S., Yakovlev A.G., Masler E.P.;
lant parasitic nematode b-TRCP";
mitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
BL; AF339101; AAK26376.1; -
terPro; IPR001680; WD40.
am; PF04000; WD40; 7.
INTS; PR00320; GPROTEINBRPT.
cdom; PD000018; WD40; 4.
AST; SM00320; WD40; 7.
OSITE; PS00678; WD_REPEATS_1; 6.
OSITE; PS50082; WD_REPEATS_2; 6.
OSITE; PS50294; WD_REPEATS_REGION; 1.
peat; WD repeat.
N TER 1
SEQUENCE 430 AA; 47916 MW; 4ABC3PF2DFE3A50B CRC64;
Match 49.4%; Score 1498; DS 5; Length 430;
Local Similarity 59.9%; Pred.No.4e-120;
es 283; Conservative 36; Mismatches 50; Indels 36; Gaps 4
193 IERMVTRDSLWRGLAERHGGQYLFPKK;-----PPDGNAP---PNSFYRALY 236
1 IERQVRNDALWAGLIRGHWKFFYCKDQATRAISQSWRPKNGEDALEFQHKFYRELY 60
237 PKLIQDIETIESNWRGRRHSIORIHCRRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNT 296
61 PKITADIRIEQWNREGRRHLRLINQCSNKGVCYCLQYDDDEKIIISGLRDNTIKIWSRPS 120
297 LECKEILLTGHTGSLVLCLOYDERRVITGSSDSTVRVWDVNTGEMLNTLIHHCVAVLHRRN 356
121 LGCQVQTLSGHTGSLVLCLOYDQNMIASGSSDQATVRIDWNTGEQKTLVHHCVAVLHRRFQ 180
357 NGMMVTCSDRSIAVMDMASPTDILRRVLVGHRAAVVVDPDDKVIIVSASGDRTIKVWN 418

Db	181	DCMLVTC	CSKORSIAVWDMA	SPDIEVRRVLVGHRAAVNVDFDHXYIVSASGDRITIKVN	240
Qy	417	TSTCEFVRTLNGH	KRGIACLQYRDRILWVSGSSDNTIRLWDIECGACLRVLGEGHELVRCI	476	
Db	241	ASTCEFVRTLNGH	KRGIACLQYRDLRILWVSGSSDYSLRWDIECGNCLRVLEGGHELVRCI	300	
Qy	477	RPDKRIVSGAYD	GKIKVWDLVAALDPRAPACTLCRLTILVEHSGRVFLQDFEQIVSSS	536	
Db	301	RFDSKRIVSGAYD	GKIMWDLAALDPMPFPNSICIKTLSQHTGRVFLQDFEQIISSS	360	
Qy	537	HDDTILWDLN	DPAA-----QAEPSPSPS	561	
Db	361	HDDTILWDLN	TCFSRAVAGAAAAAHNQVHHQAAPAAQPA	405	
RESULT 11					
ID	Q922C7	PRELIMINARY;	PRT;	252 AA.	
DT	01-DEC-2001	(TrenBMLrel. 19, Created)			
DT	01-DEC-2001	(TrenBMLrel. 19, Last sequence update)			
DT	01-MAR-2003	(TrenBMLrel. 23, Last annotation update)			
DE	Similar to f-box and WD-40 domain protein 1B (Fragment).				
DN	FBXW1B.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.				
OX	NCBI_TaxId=10090;				
[1]					
RP	SEQUENCE FROM N.A.				
RA	Strausberg R.;				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
RL	EMBL; BC008552; AAH08552.1; .				
DR	MGI; 2144023; Fbxw1b.				
DR	InterPro; IPR001680; WD40.				
DR	Pfam; PF00400; WD40; 6.				
DR	PRINTS; PR00320; GPROTEINBRPT.				
DR	ProDom; PD000018; WD40; 3.				
DR	SMART; SM00320; WD40; 5.				
DR	PROSITE; PS00678; WD_REPEATS_1; 5.				
DR	PROSITE; PS50082; WD_REPEATS_2; 6.				
DR	PROSITE; PS50294; WD_REPEATS_REGION; 1.				
KW	Repeat; WD repeat.				
FT	NON TER				
FT	1				
SEQ	SEQUENCE	252 AA; 28424 MW; F71737CBD7A9F75F	CRC64;		
Query Match	40.5%;	Score 1230;	DB 11;	Length 252;	
Best Local Similarity	92.98;	Prod No. 1.9e-97;			
Matches 234;	Conservative	7; Mismatches 11;	Indels 0;	Gaps 0;	
Qy	318	RVIITGSSDSTVRVWDVNTGEMNTLIHHCEAVLHLRFNNGMMVTCSDKORSIAVWDMA	377		
Db	1	RVIITGSSDSTVRVWDVNTGEMNTLIHHCEAVLHLRFNGLMVTCSKDRSIAVWDMA	60		
Qy	378	TDITLRRVLVGHRAAVNVDFDKKYIVSASGSDTIKWNTSTCEFVRTLNGHKGRIACIQ	437		
Db	61	TDITLRRVLVGHRAAVNVDFDKKYIVSASGORTIKWSTSTCEFVRTLNGHKGRIACIQ	120		
Qy	438	YRDLRVSGSSDNTIRLWDIECGACLRVLGEGHELVRCIFDFNKNRIVSGAYDGKIKVWDL	497		
Db	121	YRDLRVSGSSDNTIRLWDIECGACLRVLGEGHELVRCIFDFNKNRIVSGAYDGKIKVWDL	180		
Qy	498	VAAALDPRAPACTLCRLTILVEHSGRVFLQDFEQIVSSSHDDTILWDLNPPAAQAEPP	557		
Db	181	QAALDPRAPACTLCRLTILVEHSGRVFLQDFEQIISSHDDTILWDLNPPSAQNET	240		
Qy	558	RSPSRITYISR	569		
Db	241	RSPSRITYISR	252		
RESULT 12					

Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AE003480; AAL22246.1; -
DR EMBL; AY061300; AAL28848.1; -
DR EMBL; AY075401; AAL68231.1; -
DR FLYBase: FBgn0041171; ago.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF04000; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 5.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 1326 AA; 141360 MW; 3P42C873CFA3027F CRC64;

Query Match 21.8%; Score 661; DB 5; Length 1326;
Best Local Similarity 34.9%; Pred. No. 1.7e-47;
Matches 158; Conservative

QY 105 FEQSSDQVEFEVHLSIQMCHYGHSYLKPMQLQRDFITPALPARGLDHIAENILSYL 164
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 852 FORSHVERLLADRLDHCDPQSVRHMKVIFPQORDFISLLPR---ELALFVLSYL 907
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 165 DAKSLAAELVCKEWYVTSDGMLKKLIEMRVTDLSMWGLAERGWGYLFKNKPP-- 222
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 908 EPKOLLRAAQTCRSWFCLCDNLLWKCKRKA-----QILAEP-----SDRPKRG 953
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 223 -DGNAPP-NSFYRALVPKIIQDTETTESNRCGRHSIQRHCSETSKG-----VVCLOY 275
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 954 RDGNPFPIAGFWKAAAYR--QHI--JEMWR-----SRVRPKVLKHDDHVITCLQP 1003
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 276 DDQXIVSGLRDNTIKIDKNTLECKRIILTGTGSVLCLOVDERTVIITGGSDSTVRVMDVN 335
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 1004 SGNRIVSGSDNTLKWSAVNGKCLRVTGCTGVWSSQMSGNIIISGSTRTCLKVMDMD 1063
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 336 TGEMLNTLIHCEAVHLRFNNGMVTCSDRIAVIDMASPTDITLRVLVGHRAAVNV 395
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 1064 SGACVHTLQGHSTTVRCMHHLGSKVSGSRDATLRVMDIEQGSL---HYLVGHAAVRC 1120
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 396 VDFDKYIVSASGRDTIKWNVTSCFPRTLNGHKRGIAQLQVRDLRVSSGSDNTIRLW 455
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 1121 VOYDGKLIVSGAYDVYWKIWHPERQECLHTLQGHTRVYSVLPQDGLHVSGSLDTSIRVW 1180
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 456 DIEGCACLRVLGHEELVRCIGFDNKNIIVSGADVGIKWDVAALDPAPAGTLCIRTLL 515
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 1181 DVETGNCVKHTLMGHQSLSGTSMELRONILVSGNADSTVKVMDITTG-----QCQLTL 1231
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 516 V---EHSGRVFRLOQDEFQIVSSSHDDTILIWD 545
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY 1232 SGPKNKHSVATCLQFNRSFRVVTSSDDGTGVLWD 1264
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:

RESULT 13
Q9NUM6 PRELIMINARY; PRT; 553 AA.
AC Q9NUM6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ11071.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

[illegible]

OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;	
[1]	SEQUENCE FROM N.A.
RX	MEDLINE=20003061; PubMed=10531037;
RA	Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.;
RT	"A family of mammalian F-box proteins.";
RL	Curr. Biol. 9:1180-1182(1999).
RN	[2]
RP	SEQUENCE FROM N.A.
RA	Koepf D.M., Winston J.T., Harper J.W., Ellledge S.J.;
RL	Submitted (MAY-2001) to the EMBL/GenBank/DDJB databases.
DR	EMBL; AF383178; AAAG60269.1; -
DR	InterPro; IPR001810; F-box.
DR	InterPro; IPR001680; WD40.
DR	Pfam; PF00646; F-box; 1.
DR	Pfam; PF00400; WD40; 7.
DR	PRINTS; PR00320; GPROTEINBPT.
DR	PRODOM; PD000018; WD40; 2.
DR	SMART; SM00256; FBOX; 1.
DR	SMART; SM00320; WD40; 8.
DR	PROSITE; PS00181; FBOX; 1.
DR	PROSITE; PS00678; WD_REPEATS_1; 5.
DR	PROSITE; PS00082; WD_REPEATS_2; 7.
DR	PROSITE; PS02994; WD_REPEATS_REGION; 1.
KW	Repeat; WD repeat.
FT	NON_TER
SQ	SEQUENCE 561 AA; 63165 MW; B81CC1E226BD88 CRC64;
Query Match	21.1%; Score 640; DB 4; Length 561;
Best Local Similarity	30.8%; Pred No. 3.1e-46;
Matches 167; Conservative	92; Mismatches 211; Indels 72; Gaps 15;
QY	14 KPMNSEREDCNNGBPPIKIPKNSLRQTSYSCARLCNLQETVCLASTAMKTENCVAKT 73
DB	25 KLHGSEVSFSLGKKPKCV-----SEY*STTGL-----VPCSA-----TPTTFGDL 66
QY	74 KLANGSMIVPQKRLSASEYKEL--CVKYFEOWSESDOVEFYEHILISOMCHYQHGH 131
DB	67 RAANGOG-----QQRRITSVOPPTGLQEWLNKFQSWSGPEKLALDELIDSCPTQVKH 121
QY	132 INSYLKPMPLQRDPITALPARGLDHIAENTLSYLDAKSLCAABLCKEWRVRVSDGMWKK 191
DB	122 MMQVIETPOQDFISLLP---KELALYVLSPLEPKDLLQAACQCRYWRIAEADNLLWRE 177
QY	192 LIERVRTDSLNRGLAE-----RRGWGVLFNNKPPDGNAPNSFRALYPKLIQDIETI 246
DB	178 KKCE-----EGDEPUHKRK-----VIKPGFIHSPKMSAY-----IRQ--HRI 215
QY	247 ESNRCGRHSLOIRHCIRSTSETSGVYCLQYDDOKIVSGLRDNITKIWMNTLECKRILTGH 306
DB	216 DTNWARGELKS PKV-LKGHDDHVITCLOFCGNRIYVSGSDNTLVKWSAVTGKCLFLVGH 274
QY	307 TGSVLCQLQYDERVIIITGSSDSSTVRVWDVNTGEMNLTLIHCBVALHLRNPNNGMVTCSD 366
DB	275 TGGVWSSQMRDNIIGSTDTRTKVMNAFTGCICHTLYGHTSTVCRMHLHEKRVVSGSRD 334
QY	367 RSIAYVMDMASPTDITLRLVLVGHRAAVNVVDFFDKYIVSASGDRTIKWNSTCFEVRTL 426
DB	335 AYLVRVMDTETGOCL---HVLMGHVAARVCQVDDRVRVSGAVDFPVKWWDDETETCLHTL 391
QY	427 NGHKGRIACLQYRDRLVYSGSDNTIRLWIDIECGACLRVLEGHEELIVCRIFRDNKRIVSG 486
DB	392 QGHTNRVYSLPQDGIIHVYSGSLDTISRVWMDVTGNCIHTLTGHQSLSGMLKDNILVSG 451
QY	487 AYDGKIKWDLVAALDPRAPACTLCLRTLV----EHSGRVRFLOPFQIVSSSHDITILI 543
DB	452 NADSTVTKIMDKTG-----OCLQTLOGENKHQS AVTCIQFNKNFVITSSDDGTVKL 502
QY	544 WD 545
DB	WD

503 WD 504

15

96LE0 PRELIMINARY; PRT; 589 AA.

1-DEC-2001 (TremBLrel. 19, Created)

1-DEC-2001 (TremBLrel. 19, Last sequence update)

1-MAR-2003 (TremBLrel. 23, Last annotation update)

-box protein SEL10.

EL10.

omo sapiens (Human).

lkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CBI_TaxID=9606;

11

SEQUENCE FROM N.A.

1 J., Pauley A.M., Myers R.L., Shuang R., Brashler J.R., Yan R.,

hl A.E., Gurney M.E.;

SEL-10 Interacts with Presenilin 1, Facilitates Its Ubiquitination,

nd Alters A-beta Production.";

ubmitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

MBL, AY008274; RAG16640.1; ..

nterPro; IPR001810; F-box.

terPro; IPR001880; WD40.

fam; PF00646; F-box; 1.

fam; PF00400; WD40; 7.

RINTS; PRO0320; GPROTEINRPT.

coDom; PD000018; WD40; 2.

4ART; SM00256; FBOX; 1.

4ART; SM00320; WD40; 8.

ROSITE; PS50181; FBOX; 1.

ROSITE; PS00678; WD_REPEATS_1; 5.

ROSITE; PS50082; WD_REPEATS_2; 7.

ROSITE; PS50294; WD_REPEATS_REGION; 1.

peat; WD repeat.

SEQUENCE 589 AA; 66120 MW; 2AFB6E8A36E8DE CRC64;

, Match 21.1%; Score 640; DB 4; Length 589;

Local Similarity 30.8%; Pred. No. 3.3e-46;

ies 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KFMNSEREDCNNGEPKPIPEKNSLQTYNSCARLCLNQETVCLASTAMKIENTCVAKT 73

53 KLDHGEVRSFSLGKKPKV-----SEYTSITGL-----VPCSA-----TPTFGDL 94

74 KLANGTSSMIVEKQKLSAYEKEKEL--CVKYFEQWSESQDVEFVHLLISQMCHYQGH 131

95 RAANGOG-----QRRRITSVQPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVKH 149

132 INSYLKPMLODFITALPARGLDHTAENILSYLDAKSLCAELVCKEYRWVTSQGMWKK 191

150 MMQVIEPQFQDFISLLP-----KELALVLSFLEPKDLQAAQTCRYWRILAEDNLLWRE 205

192 LIERVMYRDSLWRGLAE-----RRGWGYLFKNKPPDGNAPPNSFYRALYPIQDIETI 246

206 KCCE-----EGIDEPLHIKRRK-----VIKGFHSPWKSAY-----IRQ---HRI 243

247 ESNWRCGRHSLOIRHCRSETSGVYCLQYDQKIVSGLRDNTIKIWKNTLECKRIILTGH 306

244 DTNWRREGELKSPKV-LKGHDHVIITCLQFCGNRIVSGDDNTLKVMSAVTGKCLRTLVGH 302

307 TGSVLCLQYDERVITGSDSTVRYVDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366

303 TCGWSSQMRDNIISGSDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 362

367 RSIAYVMDASPTDITLRRVLVGHRAAVNVDFDDKVIYVSGDRITIKWNTSTCEFVRTL 426

363 ATLRYWDIETGQCL---HVLGMHVAARVCQYDGRVVSQYVDFMWKVPDPTETCLHTL 419

427 NQHKRGIAQLQYRDLVSGSDNTIRLWDIECGACLRLVLEGHELVRCIRPDKRIVSG 486

Db 420 QCHTRVYSLQFDGIHVYSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 479

Qy 487 AYDGKIKVWDLVAALDPRAPAGTLCLETLV---EHSGRVPRLOQDFEQIVSSSHDDTILI 543

Db 480 NADSTVKIWDIKTG-----OCLQTLQGPKNKHOSAVTCLQFNKNFVITSSDDGTVKL 530

Qy 544 WD 545

Db 531 WD 532

Search completed: October 22, 2003, 10:25:14

Job time : 110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ein - nucleic search, using frame_plus_p2n model
October 22, 2003, 13:59:32 : Search time 4994 Seconds
(without alignments)
4661.109 Million cell updates/sec

US-09-601-168B-2
score: 3034
e: 1 MDPAAVLQKALKFMSSE.....PAAQAEPSPSRVTYIIS 569

table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 2888711 seqs, 20454813386 residues
umber of hits satisfying chosen parameters: 5777422

DB seq length: 0
DB seq length: 2000000000
ocessing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

line parameters:
frame+ p2n.model -DSV=xlp
2_1/USPTO.spool_p/US09601168/runat 22102003 100546 17778/app.query.fasta_1.711
Embl -CPWT=fastap -SUPFIX=rge -MINMATCH=0.1 -LOOSCL=0 -LOOEXT=0
bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
GN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODES=LOCAL
=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
S09601168 @CGN 1 1 3745 @runat 22102003 100546 17778 -NCPU=6 -ICPU=3
P -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
MEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
T=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

s :
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.nam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	3034	100.0	2151	6	AX019507 Sequence
2	3034	100.0	2151	6	AX057166 Sequence
3	3034	100.0	2151	6	BD131778 Human bet
4	3034	100.0	2151	9	AF129530 Homo sapi
5	3034	100.0	2151	9	Y14153 Homo sapien
6	3006	99.1	1818	9	AF101784 Homo sapi
7	3006	99.1	2308	9	BC027994 Homo sapi
8	2997	98.8	1712	10	AF099932 Mus muscu
9	2997	98.8	2081	10	BC003989 Mus muscu
10	2997	98.8	2175	10	AF081887 Mus muscu
11	2990	98.5	1710	10	AF112979 Mus muscu
12	2978	98.2	1979	10	AF110396 Mus muscu
13	2597	85.6	1671	5	M98268 African Cla
14	2400	79.1	2252	9	AB033280 Homo sapi
15	2400	79.1	4395	9	BC026213 Homo sapi
16	2384.5	78.6	1653	9	AF176022 Homo sapi
17	2384.5	78.6	2274	9	AB033281 Homo sapi
18	2384.5	78.6	4230	9	AB014596 Homo sapi
19	2380	78.4	4030	10	BC034261 Mus muscu
20	2361	77.8	2134	9	AB033279 Homo sapi
21	2360	77.8	3874	10	AB093260 Mus muscu
22	2330	76.8	1970	10	AY038079 Mus muscu
23	2250	74.2	2530	5	BC045356 Danio rer
24	2129.5	70.2	2514	3	AK114399 Ciona int
25	2128	70.1	2445	3	AB076893 Ciona int
26	2089.5	68.9	2154	3	AF032878 Drosophil
27	2089.5	68.9	2367	3	AF222924 Drosophil
28	2089.5	68.9	2534	3	AY118898 Drosophil
29	2089.5	68.9	3532	3	AF222923 Drosophil
30	1908.5	62.9	67008	9	AP000252 Homo sapi
31	1908.5	62.9	100000	9	AP000212 Homo sapi
32	1908.5	62.9	149298	9	AP000031 Homo sapi
33	1908.5	62.9	340000	9	AP001711 Homo sapi
34	1685	55.5	173634	3	AC009344 Drosophil
35	1685	55.5	235707	3	AE003733 Drosophil
36	1683.5	55.5	292639	2	AC112602 Rattus no
37	1653.5	54.5	40472	2	AC014085 Drosophil
38	1638.5	54.0	2700	3	AF275253 Caenorhab
39	1503.5	49.6	17344	3	U28730 Caenorhabdi
40	1498	49.4	1443	3	AF339101 Heteroder
41	1230	40.5	2983	10	BC008552 Mus muscu
42	1189	39.2	1827	5	U63922 Xenopus lae
43	1180	38.9	934	5	XLU63921 Xenopus lae
44	1163	38.3	657	6	AX505273 Sequence
45	1075	35.4	1217	3	AK115763 Ciona int

ALIGNMENTS

7	AX019507	2151 bp	DNA	linear	PAT 07-SEP-2000	
ION	Sequence 1 from Patent WO9938969.					
ON	AX019507					
S	AX019507.1	GI:10043427				
ISM	synthetic construct					
	synthetic construct					
	artificial sequences.					
CE	1					
RS	Arenzana, S.F., Concordet, J.P., Kroll, M., Durand, H., Benarous, R. and Margottin, F.					
	Protein humaine beta-trop.					
AL	Patent: WO 9938969-A1 05-AUG-1999; ARENZANA SEISDEDOS FERNANDO (FR); CONCORDET JEAN PAUL (FR); INST NAT SANTE RECH MED (FR); KROLL MATHIAS (FR); DURAND HERVE (FR); BENAROUS RICHARD (FR); MARGOTTIN FLORENCE (FR); PASTEUR INSTITUT (FR)					
S	Location/Qualifiers					
urce	1. .2151					
	/organism="synthetic construct"					
	/mol type="genomic DNA"					
	/db xref="taxon:32630"					
	/note="ADNC codant pour la proteine humaine BTRCP"					
S	70..1779					
	/note="unnamed protein product"					
	/codon_start=1					
	/transl_table=11					
	/protein_id="CAC07740.1"					
	/db xref="GI:10043428"					
	/translation="NDPAVLQEKALFPMSSEREDNCNCEPPPKIIPKNSLRQTV NSCARCLNQEVCLASTAKWNTKCVAKTKLANGTSSMIVPKQKLSASVKEKELCV KFEQSESDQEFVHLLISQCHVQGHINSYLPKMLQRFITLARGDLHIAENI LSYLDASLCALAEVCKEYRVTSQDGLWKKLIERWTDLSRLGRABRRGWQGLFK NRPDQAPNPFYALYFKIIQDIETIESNRWCRHSLSQRIHCRSETSKGVYCLQYD DQIVSGLRDNTIKIWDKNTLCKRILITGHTGSLCLQYDERVITGSSDSFVRWQV NTGEMLNTLIHCEAVLHLRFNMGMVTCSDRSIAVMDMSPDTITLRRVLVGHRAA VNVVDKKIVYSASGDRITIKYNTSTCFEFTLNGHKRGACILOYRDLRVYSSGSDN TIELWIEGACLRVLECHEELVRCIRFPNKRIVSGAYDGGIKIWDWLVAAALDPRAPAG TLCLRLVSHSGVRFRLQDFQDEFQVSSHDDTLILWDFLNDPAQAQEPSPSRITY ISR"					
UNT	628 a	467 c	513 g	543 t		
nt Scores:						
S:	3.75e-274	Length:	2151			
Similarity:	3034.00	Matches:	569			
cal Similarity:	100.00%	Conservative:	0			
atch:	100.00%	Mismatches:	0			
	100.00%	Indels:	0			
	6	Gaps:	0			
01-168B-2 (1-569) x AX019507 (1-2151)						
1	MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu	20				
70	ATGACCCGCCCGAGGGGGTGTGCAAGAGAGAGGCACTCAAGTTTATGATTAATCCTCAGAG	129				
21	ArgGluAspCysAsnAndGlyGluProProArgLysIleProGluLysAsnSerLeu	40				
130	AGAGAAGACTGTATATATGCGCAACCCCTTAGGAAGATAATACGAGAGAAATTCACCT	189				
41	ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla	60				
190	AGACAGACATACACACAGCTGGCCAGACCTCTGCTTAACCAAGAACACATGATTGATGA	249				
61	SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer	80				
250	AGCACTGCTATGAGACTGAGAAATTTGTGGCCAAACAAACTTGCCCAATGGCACTTCC	309				
81	SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu	100				
310	AGTATGATTTGTGCCCAAGCAACGGAACCTCTAGCAAGCTATGAAAGAAAGAAAGCACTG	369				
101	CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu	120				
370	TGTGTCAATACTTTTGGAGAGTGGTTCAGATCAAGTGAATTTTGGAAATCTTT	429				
121	IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu	140				
430	ATATCCCAATGTGTTCATACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG	489				
141	GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle	160				
490	CAGAGAGATTTCACTAATCTCTGCCAGCTCGGGATTTGGATCATATCGTGAGAACATT	549				
161	LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr	180				
550	CTGTATACCTGGATGCCAAATCACTATGTGTCTGTAACCTGTGTGCAAGGAATGGTAC	609				
181	ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp	200				
610	CGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCGAGAGAAATGGTTCAGGACAGAT	669				
201	SerLeuTrpArgGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLys	220				
670	TCTCTGTGGAGAGCCCTGCGCAGAACGAGAGATGGGACAGTATTTATTTCAAAAACAAA	729				
221	ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle	240				
730	CCTCTGACGGGAATGCTCTCTCCCACTCTTTTATAGAGCACTTTATCTCTAAAATTTATA	789				
241	GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle	260				
790	CAAGACATTCAGACATAGAACTAATTCGAGATGTGGAAGACATAGTTTACAGAGAAAT	849				
261	HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle	280				
850	CACCTGCGAAGTGAACAAAGAGAGTGTCTGTTTACAGATGATGATCAGAAAAATA	909				
281	ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys	300				
910	GTAAGCGCCCTTCGAGACAAACAACTCAAGATCTGGATTAATAAACACATTTGGAATGCAAG	969				
301	ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle	320				
970	CGAATTCACAGGCCATACAGGTTCCAGTCTCTGCTCCAGTATGATGAGAGAGTGATC	1029				
321	IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu	340				
1030	ATAACAGGATCATCGGATTCACCGGTCCAGAGTGGGATGTAAATACAGGTGAAATGCTTA	1089				
341	AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet	360				
1090	AACAGTTGATTCAACCATTTGTAAGCAGTTCTGCACCTTCCGTTTCAATAATGCGCATGATG	1149				
361	ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle	380				
1150	GTGACCTGCTCCAAAGATCGTTCCATTTGTTATGGATATGCGCTCCCAACTGCACATT	1209				
381	ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp	400				
1210	ACCTTCGGAGGGTGTGTTGCGACACCGAGCTGCTGTCATATGTTGTAGACTTTGATGAC	1269				
401	LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys	420				
1270	AAGTACATTTCTCTCATCTGGGGATAGAACTATATAAGGTATGGAACACAAGTACTTGT	1329				
421	GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp	440				
1330	GAATTTGTAGGACCTTAAATGGACAAACGAGGCAATGCGCTGTTTGCAGTACAGGGAC	1389				
441	ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly	460				
1390	AGGCTGTTGTAGTGGCTCATCTGCAACACTATCAGATTATGGACATAGANTGTGT	1449				

461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACGAGTGTAGAGGCCATGAGGAATTGGTGGTTGTATTCGATTGATAAC 1509
 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
 1510 AAGAGGATAGTCAGTGGGGCTATGATGGAAAAATTAAGATGGGGATCTTGTGGCTGCT 1569
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGGACCCCGTCTCTCTGAGGACACCTCTGTCTACGGACCTTGTGGAGCATTCGGGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTTCGACTACAGTTTGTATGATTCAGATTGTCCAGATTGTCCAGTATTCACATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
 1690 ATCTCATCTGGGACTTCTTAATGATCCAGTGGCCAGCTGACCCCGCTTCCCT 1749
 561 SerArgThrTyrThrTyrIleSerArg 569
 1750 TCTCGAACATACACCTACATCTCCAGA 1776
 2
 6
 AX057166 2151 bp DNA linear PAT 17-JAN-2001
 ION Sequence 42 from Patent WO0075184.
 ON AX057166
 AX057166.1 GI:12309979
 S Homo sapiens (human)
 ISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CE 1
 RS Zhang, H., Tsvetkov, L.M. and Kondo, T.
 AL Modulation of protein levels using the scf complex
 Patent: WO 0075184-A 42/14-DEC-2000;
 YALE UNIVERSITY (US)
 S Location/Qualifiers
 1..2151
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 UNT 628 a 467 c 513 g 543 t
 nt Scores:
 O.: 3.75e-274 Length: 2151
 Similarity: 3034.00 Matches: 569
 Conservativity: 0
 cal Similarity: 100.00% Mismatches: 0
 Indels: 0
 Gaps: 0
 01-168B-2 (1-569) x AX057166 (1-2151)
 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 70 ATGGACCCGCCGAGGGGTGCTGCAAGAGAGAGCACTCAAGTTTATGAATTCCTCAGAG 129
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleLeuProGluLysAsnSerLeu 40
 130 AGAGAAGACTGTAATATATGCGAACCCCTTAGGAAGATAATACAGAGAAGAAATTCACCT 189
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlnThrValCysLeuAla 60
 190 AGACAGACATACACAGCTGTGCGACACTCTGCTTAAACCAAGAAACAGATGTTTAGCA 249
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 250 AGCACTGCTATGAAGACTGAGAAATTGTGTGGCCAAACAAACTTGCCCAATGGCACTTCC 309

QY 81 SerWetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 310 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAGACTATGAAAGGAAAAGGAACTG 369
 QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 370 TGTGTCAATACATTGAGCAGTGTGTGAGAGTCAAGTCAAGTGGAAATTTGTGGAACATCTT 429
 QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 430 ATAATCCCAATGTGTCAATACCAACTGCGCAATAAACTCGTATCTTAAACCTATGTTG 489
 QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 490 CAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGATGGATCATATCGTCGAGAACATT 549
 QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 550 CTGTATACCTGATGCGCAATCACTATGTGCTGCTGAACTTGTGTGCAAGGAATGGTAC 609
 QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGAAATGGTCAGGACAGAT 669
 QY 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
 Db 670 TCTCTGTGGAGAGGCGCTGGCAGAACGAAGAGGATGGGACAGTATTTATTCAAAAACAAA 729
 QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 Db 730 CCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCTTAAAAATTATA 789
 QY 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 Db 790 CAAGACATTGAGACAATAGAAATCTAATTGGAGATGTGGAGACATAGTTTACAGAAATT 849
 QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 Db 850 CACTGCCGAAGTGAACCAAGCAAGAGGAGTTTACTGTTTACAGTATGATCAGAAAAATA 909
 QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 Db 910 GTAAGCGCGCTTCGAGACAACACAATCAAGATCTGGGATAAAACACATTTGGAATCAAG 969
 QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 Db 970 CGAATTCACAGGCCATACAGGTTCAAGTCTCTCTCTCCAGTATGATGAGAGAGTGATC 1029
 QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 Db 1030 ATACAGGATCATCGGATTCACGGTCAGAGTGTGGATGTAATAACAGGTGAATTCCTA 1089
 QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 Db 1090 AACACGTTGATTACCAATTTGGAAGCAGTTTCGACTTGGCTTCAATATGCGATGATG 1149
 QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 Db 1150 GTGACCTGCTCCAAAGATGCTTCATTTGCTGTATGGATATGCGCTCCCACTGACATT 1209
 QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 Db 1210 ACCCTCCGAGGGTGTGGTCCGACACCGAGCTGCTCTCAATGTTCTAGACTTTGTATGAC 1269
 QY 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 Db 1270 AAGTACATTGTTTCTGCACTCTGGGATAGAACTATTAAGGTATGGAACACAACTACTTGT 1329
 QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 Db 1330 GAATTTGTAGGACCTTAATGTGACACAACAGGAGCATTCGCTGTTTGGAGTACAGGAC 1389

441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1390 AGGCTGGTAGTGGCTCATCTGACAACTATCAGATTATGGGACATAGAAATGGT 1449
461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
1450 GCATGTTTACAGTGTAGAGGCCATGAGAAATGGTGTTGTTATTCGATTGTAAAC 1509
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
1510 AAGAGGATAGTCAGTGGGCTATGATGGAAAAATTAAGTGTGGGATCTTGTGGCTGCT 1569
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1570 TTGGACCCCGTCTCTCGAGGACACTCTGTCTACGGACCCCTGTGGAGCATTCGGGA 1629
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
1630 AGAGTTTTCGATACAGTTTGTAGTATCCAGATTGTGAGTATTCATGATGACACA 1689
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
1690 ATCTCATCTGGGACTTCTTAATGATCCAGCTGCCAAGCTCAACCCCGTTCCTCT 1749
561 SerArgThrTyrThrTyrIleSerArg 569
1750 TCTCGAACATACACCTACATCTCCAGA 1776

BD131778 2151 bp DNA linear PAT 18-SEP-2002
Human beta TrCP protein.
BD131778
BD131778.1 GI:23226723
JP 2002501746-A/1.
synthetic construct
synthetic construct
artificial sequences.
1 (bases 1 to 2151)
Benarous, R., Margottin, F., Durand, H., Seisdedos, F.A., Kroll, M. and
Concordet, J.P.
Human beta TrCP protein
Patent: JP 2002501746-A 10/22-JAN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, INSTITUT
PASTEUR
OS Artificial Sequence
PN JP 2002501746-A/1
PD 22-JAN-2002
PF 29-JAN-1999 JP 2000529429
PR 30-JAN-1998 FR 98/01100, 09-DEC-1998 FR 98/15545 PI
RICHARD BENAROUS, FLORENCE MARGOTTIN, HERVE DURAND, PI FERNANDO
ARENZANA SEISDEDOS, MATHIAS KROLL, JEAN PAUL CONCORDET PC
C12N15/00, A01K67/027, A61K38/00, A61K45/00, A61P29/00, A61P31/12, PC
A61P35/00,
PC
C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12Q1/68, C12P21/ PC
02,
PC C12N15/00, A61K37/02, C12N5/00
CC Description of the artificial sequence :
ADNC coding for human
CC (TrCP
CC protein
CC key
FH Location/Qualifiers
FT CDS (70)..(1776).
Location/Qualifiers
1..2151
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
628 a 467 c 513 g 543 t

nt Scores: *

Pred. No.: 3.75e-274 Length: 2151
Score: 3034.00 Matches: 569
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-601-168B-2 (1-569) x BD131778 (1-2151)
QY 1 MetAspProAlaGluAlaValLeuGlnGlnLysAlaLeuLysPheMetAsnSerSerGlu 20
DB 70 ATGACCCCGCGGAGGGCTGCTGCAAGAGAGAGCACTCAAGTTTATGAATTTCTCAGAG 129
QY 21 ArgGluAspCysAsnAsnGlyGluProArgIleLysIleProGluLysAsnSerLeu 40
DB 130 AGAAGAGACTGTAATAATGGCGAACCCCTAGGAAGATATATACCAGAGAGAAATTCATT 189
QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
DB 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCA 249
QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
DB 250 AGCACTGCTATGAAGACTGAGAAITGTGCCCAAAACAAACTTGCCATGGCACTTCC 309
QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
DB 310 AGTATGATTGTGCCAAGCAACGGAACTCTCAGCAAGCTATGAAAGGAAAGGAACTG 369
QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
DB 370 TGTGTCAAAATCTTTGAGCAGTGTGTGAGAGTCAGATCAAGTGGAAATTTGTGGAACATCTT 429
QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
DB 430 ATATCCCAATGTGTCTATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 489
QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
DB 490 CAGAGAGATTTTCATAACTGTCTGCCAGCTCGGGGATTGGATCATATCGCTGAGAACATT 549
QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaIleGluLeuValCysLysGluTrpTyr 180
DB 550 CTGTACATACCTGGATGCGAAATCACTATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
DB 610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGAAATGGTCAGGACAGAT 669
QY 201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
DB 670 TCTCTGTGGAGAGGCTGGCAGAACGAGAGGATGGGACAGTATTATTATTTCAAAAACAAA 729
QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
DB 730 CCTCTGTACGGGAATGCTCTCTCCCACTCTTTTATAGACACTTTATCTCTAAATATA 789
QY 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
DB 790 CAAGACATTGAGACAATAGAAATCAATTTGGAGATGTGGAAGACATAGTTTACAGAGAAAT 849
QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
DB 850 CACTGCCGAGTGAACAAGCAAGCAAGGAGTGTACTGTTTACAGTATGATGATCAGAAAATA 909
QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
DB 910 GTAAGCGGCTTTCAGACAACAATCAAGATCTGGGATATAAAACACATTTGGAATGCAAG 969
QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
DB 970 CGAATTTCTCAGAGGCCATACAGGTTTCAGTCTCTGCTCTCCAGTATGATGAGAGAGTATC 1029

321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 1030 ATAACAGATCATCGGATTCACCGCTCAGATGTGGGATGTAATACAGGTCAAATGCTA 1089
 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 1090 AACAGTTGATTCACCATTTGAGACGAGTCTGCACCTTGCCTTTCAATAATGGCATGATG 1149
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1150 GTGACCTGCTCCAAAGATCGTTCATTCGTGTATGGGATATGGCTCCCAACTGACATT 1209
 381 ThrLeuArgValLeuValGlyHisArgAlaValAsnValAspPheAspAsp 400
 1210 ACCCTCCGAGGAGTCTGCTCGGACACCGAGCTGCTGCAATGTGTAGACTTTGATGAC 1269
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1270 AAGTACATTTGCTGCACTCGGGATAGAACTATAAGGTATGGAACACAGTACTTGT 1329
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 1330 GAATTTGTAAGACCTTAATGGACACAAACGAGGCAITTCCTGTTTCAGTACAGGAC 1389
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1390 AGGTGGTAGTGGTCTCATCTGACAACTATCAGATATGCGGACATAGATGTGGT 1449
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACAGTGTTAGAGGCCATGAGGAATGTGGCTGTATTCGATTCGATTAAC 1509
 481 LysArgIleValSerGlyValTyrAspGlyLysIleLysValTrpAspLeuValAla 500
 1510 AAGAGGATAGTCAGTGGGGCTATGATGAAATAATTAAGTGTGGATCTTGTGGCTGCT 1569
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGGACCCCGTCTCTGCGAGGACACTCTGTCTACGGACCTCTGTGGACATTCGCGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTCGACTACAGTTGATGAATTCAGATGTCAGATGTCACATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1690 ATCCTCATCTCGGACTTCTTAATGATCCAGCTGCCAAGCTGAACCCCGCTCCCT 1749
 561 SerArgThrTyrThrTyrIleSerArg 569
 1750 TCTCGAATACACCTACATCTCCAGA 1776

AF129530 2151 bp mRNA linear PRI 31-OCT-1999
 ON Homo sapiens chromosome 10 P-box protein Fbw1A (Fbw1A) mRNA,
 complete cds.

AF129530
 AF129530.1 GI:6164609

SM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2151)
 Cenciarelli, C., Chiau, D.S., Guardavaccaro, D., Parks, W., Vidal, M.
 and Pagano, M.

Identification of a family of human P-box proteins
 Curr. Biol. 9 (20), 1177-1179 (1999)

FE 2003060

FE 10531035

FE 2 (bases 1 to 2151)

FE Chiau, D.S. and Pagano, M.

S

TITLE Direct Submission
 JOURNAL Submitted (19-FEB-1999) Pathology, NYU Medical Center, 550 First
 Ave. MSB 548, New York, NY 10016, USA

FEATURES Location/Qualifiers

source

1..2151
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /map="10q24"
 1..2151
 /gene="FBW1A"
 70..1779
 /gene="FBW1A"
 /function="regulates the stability of IkappaB alpha- and
 beta-catenin"

/note="contains one P-box domain and seven WD40 domains;
 forms an SCF with Skp1 and Cull1"

/codon_start=1
 /product="P-box protein Fbw1A"
 /protein_id="AAF04464.1"
 /db_xref="GI:6164610"
 /translation="NDPAEAVLQEKALFKMNSREDGNCGEPPPKIIPKNSLRQTY
 NSCARLCLNQETVCLASTAMKTENCVCATKLANGTSSMIVFKQKLSASYEKELCV
 KYFEQWSSDQVEFVEHLISQMHCHYQHGHIINSYLPMLQDFITALPARGLDHIAENI
 LSYLDAKLAELVKEMRYRTSDGLWKKLIERMVETDSLRLGRLAERGMGQYLFK
 NKPPDGNAPPSFYRALYFKIITDIETIESNWRGSHSLQRIHCRSETSKGVYCLQYD
 DOKIVSLRDLTIKIWKDNTLECKRLTGHGTVLCLOYDERVITIGSSDSTVRVMDV
 NTGEMLTLIHHCBAVLHLPNNGMMVTCSDKRSIATVMDASPTDITLRLVLVGHRAA
 VNVVDFDKYIVSASGDRITIKVNTSTCEFYTLNGHKGRIACLOYRLVLYVSSSDN
 TLRMLDIEGACLRVLEGHLEELVRCIRFDNKRIVSGAYDVGKIKVMDLVAAALDPAPAG
 TILCLRTLVEHSGRVFRLOFDEFQIVSSSHDPTILIMDFNDPAAQAEPPSPSRITYY
 ISR"

BASE COUNT 628 a 467 c 513 g 543 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,75e-274 Length: 2151
 Score: 3034.00 Matches: 569
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-601-168B-2 (1-569) x AF129530 (1-2151)

QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 70 ATGGACCCGCGGAGCGGTGCTCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
 QY 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
 Db 130 AGAGAAGACTGTAATAATGGCAACCCCTAGGAAGATAATACCAGAGAAGAAATTCACCT 189
 QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 190 AGACAGACATACACAGCTGTGCCAGACTCTCTTAAACCAAGAACAGATATGTTTAGCA 249
 QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 250 AGCAGCTGCTATGAGACTGAGAAATGTGTGGGCAAAACAAACTTGGCCATGGCACTCC 309
 QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 310 AGTATGATTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAAGGAAAAGGAACGTG 369
 QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 370 TGTGTCAAATACTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGAATTTGTGGAACATCTT 429
 QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 430 ATATCCCAANTGTGTCAATACCAACATGGGCAATAAACTCGTATCTTAAACCTATGTTG 489

```

141  GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
142  |||||
143  |||||
144  |||||
145  |||||
146  |||||
147  |||||
148  |||||
149  |||||
150  |||||
151  |||||
152  |||||
153  |||||
154  |||||
155  |||||
156  |||||
157  |||||
158  |||||
159  |||||
160  |||||
161  LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysGlyGluTyrTyr 180
162  |||||
163  |||||
164  |||||
165  |||||
166  |||||
167  |||||
168  |||||
169  |||||
170  |||||
171  |||||
172  |||||
173  |||||
174  |||||
175  |||||
176  |||||
177  |||||
178  |||||
179  |||||
180  |||||
181  ArgValThrSerAspGlyMetLeuTyrLysLeuIleGluArgMetValArgThrAsp 200
182  |||||
183  |||||
184  |||||
185  |||||
186  |||||
187  |||||
188  |||||
189  |||||
190  |||||
191  |||||
192  |||||
193  |||||
194  |||||
195  |||||
196  |||||
197  |||||
198  |||||
199  |||||
200  |||||
201  SerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
202  |||||
203  |||||
204  |||||
205  |||||
206  |||||
207  |||||
208  |||||
209  |||||
210  |||||
211  |||||
212  |||||
213  |||||
214  |||||
215  |||||
216  |||||
217  |||||
218  |||||
219  |||||
220  |||||
221  ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
222  |||||
223  |||||
224  |||||
225  |||||
226  |||||
227  |||||
228  |||||
229  |||||
230  |||||
231  |||||
232  |||||
233  |||||
234  |||||
235  |||||
236  |||||
237  |||||
238  |||||
239  |||||
240  |||||
241  GlnAspIleGluThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIle 260
242  |||||
243  |||||
244  |||||
245  |||||
246  |||||
247  |||||
248  |||||
249  |||||
250  |||||
251  |||||
252  |||||
253  |||||
254  |||||
255  |||||
256  |||||
257  |||||
258  |||||
259  |||||
260  |||||
261  HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
262  |||||
263  |||||
264  |||||
265  |||||
266  |||||
267  |||||
268  |||||
269  |||||
270  |||||
271  |||||
272  |||||
273  |||||
274  |||||
275  |||||
276  |||||
277  |||||
278  |||||
279  |||||
280  |||||
281  ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
282  |||||
283  |||||
284  |||||
285  |||||
286  |||||
287  |||||
288  |||||
289  |||||
290  |||||
291  |||||
292  |||||
293  |||||
294  |||||
295  |||||
296  |||||
297  |||||
298  |||||
299  |||||
300  |||||
301  ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
302  |||||
303  |||||
304  |||||
305  |||||
306  |||||
307  |||||
308  |||||
309  |||||
310  |||||
311  |||||
312  |||||
313  |||||
314  |||||
315  |||||
316  |||||
317  |||||
318  |||||
319  |||||
320  |||||
321  IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
322  |||||
323  |||||
324  |||||
325  |||||
326  |||||
327  |||||
328  |||||
329  |||||
330  |||||
331  |||||
332  |||||
333  |||||
334  |||||
335  |||||
336  |||||
337  |||||
338  |||||
339  |||||
340  |||||
341  AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnGlyMetMet 360
342  |||||
343  |||||
344  |||||
345  |||||
346  |||||
347  |||||
348  |||||
349  |||||
350  |||||
351  |||||
352  |||||
353  |||||
354  |||||
355  |||||
356  |||||
357  |||||
358  |||||
359  |||||
360  |||||
361  ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
362  |||||
363  |||||
364  |||||
365  |||||
366  |||||
367  |||||
368  |||||
369  |||||
370  |||||
371  |||||
372  |||||
373  |||||
374  |||||
375  |||||
376  |||||
377  |||||
378  |||||
379  |||||
380  |||||
381  ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
382  |||||
383  |||||
384  |||||
385  |||||
386  |||||
387  |||||
388  |||||
389  |||||
390  |||||
391  |||||
392  |||||
393  |||||
394  |||||
395  |||||
396  |||||
397  |||||
398  |||||
399  |||||
400  |||||
401  LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420
402  |||||
403  |||||
404  |||||
405  |||||
406  |||||
407  |||||
408  |||||
409  |||||
410  |||||
411  |||||
412  |||||
413  |||||
414  |||||
415  |||||
416  |||||
417  |||||
418  |||||
419  |||||
420  |||||
421  GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
422  |||||
423  |||||
424  |||||
425  |||||
426  |||||
427  |||||
428  |||||
429  |||||
430  |||||
431  |||||
432  |||||
433  |||||
434  |||||
435  |||||
436  |||||
437  |||||
438  |||||
439  |||||
440  |||||
441  ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460
442  |||||
443  |||||
444  |||||
445  |||||
446  |||||
447  |||||
448  |||||
449  |||||
450  |||||
451  |||||
452  |||||
453  |||||
454  |||||
455  |||||
456  |||||
457  |||||
458  |||||
459  |||||
460  |||||
461  AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
462  |||||
463  |||||
464  |||||
465  |||||
466  |||||
467  |||||
468  |||||
469  |||||
470  |||||
471  |||||
472  |||||
473  |||||
474  |||||
475  |||||
476  |||||
477  |||||
478  |||||
479  |||||
480  |||||
481  LysArgIleValSerGlyValTyrAspGlyLysIleLysValTyrAspLeuValAlaAla 500
482  |||||
483  |||||
484  |||||
485  |||||
486  |||||
487  |||||
488  |||||
489  |||||
490  |||||
491  |||||
492  |||||
493  |||||
494  |||||
495  |||||
496  |||||
497  |||||
498  |||||
499  |||||
500  |||||
501  LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
502  |||||
503  |||||
504  |||||
505  |||||
506  |||||
507  |||||
508  |||||
509  |||||
510  |||||
511  |||||
512  |||||
513  |||||
514  |||||
515  |||||
516  |||||
517  |||||
518  |||||
519  |||||
520  |||||
521  ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
522  |||||
523  |||||
524  |||||
525  |||||
526  |||||
527  |||||
528  |||||
529  |||||
530  |||||
531  |||||
532  |||||
533  |||||
534  |||||
535  |||||
536  |||||
537  |||||
538  |||||
539  |||||
540  |||||
541  IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
542  |||||
543  |||||
544  |||||
545  |||||
546  |||||
547  |||||
548  |||||
549  |||||
550  |||||
551  |||||
552  |||||
553  |||||
554  |||||
555  |||||
556  |||||
557  |||||
558  |||||
559  |||||
560  |||||
561  SerArgThrTyrThrTyrIleSerArg 569
562  |||||
563  |||||
564  |||||
565  |||||
566  |||||
567  |||||
568  |||||
569  |||||
570  |||||
571  |||||
572  |||||
573  |||||
574  |||||
575  |||||
576  |||||
577  |||||
578  |||||
579  |||||
580  |||||
581  |||||
582  |||||
583  |||||
584  |||||
585  |||||
586  |||||
587  |||||
588  |||||
589  |||||
590  |||||
591  |||||
592  |||||
593  |||||
594  |||||
595  |||||
596  |||||
597  |||||
598  |||||
599  |||||
600  |||||
601  |||||
602  |||||
603  |||||
604  |||||
605  |||||
606  |||||
607  |||||
608  |||||
609  |||||
610  |||||
611  |||||
612  |||||
613  |||||
614  |||||
615  |||||
616  |||||
617  |||||
618  |||||
619  |||||
620  |||||
621  |||||
622  |||||
623  |||||
624  |||||
625  |||||
626  |||||
627  |||||
628  |||||
629  |||||
630  |||||
631  |||||
632  |||||
633  |||||
634  |||||
635  |||||
636  |||||
637  |||||
638  |||||
639  |||||
640  |||||
641  |||||
642  |||||
643  |||||
644  |||||
645  |||||
646  |||||
647  |||||
648  |||||
649  |||||
650  |||||
651  |||||
652  |||||
653  |||||
654  |||||
655  |||||
656  |||||
657  |||||
658  |||||
659  |||||
660  |||||
661  |||||
662  |||||
663  |||||
664  |||||
665  |||||
666  |||||
667  |||||
668  |||||
669  |||||
670  |||||
671  |||||
672  |||||
673  |||||
674  |||||
675  |||||
676  |||||
677  |||||
678  |||||
679  |||||
680  |||||
681  |||||
682  |||||
683  |||||
684  |||||
685  |||||
686  |||||
687  |||||
688  |||||
689  |||||
690  |||||
691  |||||
692  |||||
693  |||||
694  |||||
695  |||||
696  |||||
697  |||||
698  |||||
699  |||||
700  |||||
701  |||||
702  |||||
703  |||||
704  |||||
705  |||||
706  |||||
707  |||||
708  |||||
709  |||||
710  |||||
711  |||||
712  |||||
713  |||||
714  |||||
715  |||||
716  |||||
717  |||||
718  |||||
719  |||||
720  |||||
721  |||||
722  |||||
723  |||||
724  |||||
725  |||||
726  |||||
727  |||||
728  |||||
729  |||||
730  |||||
731  |||||
732  |||||
733  |||||
734  |||||
735  |||||
736  |||||
737  |||||
738  |||||
739  |||||
740  |||||
741  |||||
742  |||||
743  |||||
744  |||||
745  |||||
746  |||||
747  |||||
748  |||||
749  |||||
750  |||||
751  |||||
752  |||||
753  |||||
754  |||||
755  |||||
756  |||||
757  |||||
758  |||||
759  |||||
760  |||||
761  |||||
762  |||||
763  |||||
764  |||||
765  |||||
766  |||||
767  |||||
768  |||||
769  |||||
770  |||||
771  |||||
772  |||||
773  |||||
774  |||||
775  |||||
776  |||||
777  |||||
778  |||||
779  |||||
780  |||||
781  |||||
782  |||||
783  |||||
784  |||||
785  |||||
786  |||||
787  |||||
788  |||||
789  |||||
790  |||||
791  |||||
792  |||||
793  |||||
794  |||||
795  |||||
796  |||||
797  |||||
798  |||||
799  |||||
800  |||||
801  |||||
802  |||||
803  |||||
804  |||||
805  |||||
806  |||||
807  |||||
808  |||||
809  |||||
810  |||||
811  |||||
812  |||||
813  |||||
814  |||||
815  |||||
816  |||||
817  |||||
818  |||||
819  |||||
820  |||||
821  |||||
822  |||||
823  |||||
824  |||||
825  |||||
826  |||||
827  |||||
828  |||||
829  |||||
830  |||||
831  |||||
832  |||||
833  |||||
834  |||||
835  |||||
836  |||||
837  |||||
838  |||||
839  |||||
840  |||||
841  |||||
842  |||||
843  |||||
844  |||||
845  |||||
846  |||||
847  |||||
848  |||||
849  |||||
850  |||||
851  |||||
852  |||||
853  |||||
854  |||||
855  |||||
856  |||||
857  |||||
858  |||||
859  |||||
860  |||||
861  |||||
862  |||||
863  |||||
864  |||||
865  |||||
866  |||||
867  |||||
868  |||||
869  |||||
870  |||||
871  |||||
872  |||||
873  |||||
874  |||||
875  |||||
876  |||||
877  |||||
878  |||||
879  |||||
880  |||||
881  |||||
882  |||||
883  |||||
884  |||||
885  |||||
886  |||||
887  |||||
888  |||||
889  |||||
890  |||||
891  |||||
892  |||||
893  |||||
894  |||||
895  |||||
896  |||||
897  |||||
898  |||||
899  |||||
900  |||||
901  |||||
902  |||||
903  |||||
904  |||||
905  |||||
906  |||||
907  |||||
908  |||||
909  |||||
910  |||||
911  |||||
912  |||||
913  |||||
914  |||||
915  |||||
916  |||||
917  |||||
918  |||||
919  |||||
920  |||||
921  |||||
922  |||||
923  |||||
924  |||||
925  |||||
926  |||||
927  |||||
928  |||||
929  |||||
930  |||||
931  |||||
932  |||||
933  |||||
934  |||||
935  |||||
936  |||||
937  |||||
938  |||||
939  |||||
940  |||||
941  |||||
942  |||||
943  |||||
944  |||||
945  |||||
946  |||||
947  |||||
948  |||||
949  |||||
950  |||||
951  |||||
952  |||||
953  |||||
954  |||||
955  |||||
956  |||||
957  |||||
958  |||||
959  |||||
960  |||||
961  |||||
962  |||||
963  |||||
964  |||||
965  |||||
966  |||||
967  |||||
968  |||||
969  |||||
970  |||||
971  |||||
972  |||||
973  |||||
974  |||||
975  |||||
976  |||||
977  |||||
978  |||||
979  |||||
980  |||||
981  |||||
982  |||||
983  |||||
984  |||||
985  |||||
986  |||||
987  |||||
988  |||||
989  |||||
990  |||||
991  |||||
992  |||||
993  |||||
994  |||||
995  |||||
996  |||||
997  |||||
998  |||||
999  |||||
1000  |||||

```

/rpt unit=516. .544
UNT 628 a 467 c 513 g 543 t

nt Scores:

3.75e-274 Length: 2151
3034.00 Matches: 569
Similarity: 100.00% Conservative: 0
cal Similarity: 100.00% Mismatches: 0
atch: 100.00% Indels: 0
9 Gaps: 0

01-168B-2 (1-569) x HSBTRCP (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
70 ATGGACCCCGCGCGCGTGTGCAAGAGAGAGGACCTCAAGTTTATGAATTCCTCAGAG 129
21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
130 AGAGAGAGACTGTAATAATGGCGAACCCCTAGAGAGATAATACAGAGAGAGATTCATT 189
41 ArgGlnThrTyAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTTAAACCAAGAAACAGTAGTTTACGA 249
61 SerThrAlaMetLysThrGluAsnGlyValAlaLysThrLysLeuAlaAsnGlyThrSer 80
250 AGACATGCTATGAAGACTGAGAAATGTGTGGCCAAAAACAACCTTGCCAAATGGCACITCC 309
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrgLysGluLysGluLeu 100
310 AGTATGATTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAAGGAACTG 369
101 CysValLysTyPheGluGlnThrPheSerGluSerAspGlnValGluPheValGluHisLeu 120
370 TGTGTCAAAATCTTTGAGCAGTGGTCAGAGTCAGATCAGATCAAGTGAATTTGTGGAACATCTT 429
121 IleSerGlnMetCysHisTyrgLysGlnHisGlyHisIleAsnSerTyrgLysProMetLeu 140
430 ATATCCAAATGTGTATTCACCAATGGGCACATTAACCTGATCTTAACCTATGTTG 489
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
490 CAGAGAGATTCATAACTGCTGCAGCTCGGGATGGATCATATCGCTGAGAAACATT 549
161 LeuSerTyrgLysAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluThrTyrg 180
550 CTGTCAATCTGGATGCCAAATCATTATGTGTCTGTAACCTGTGTGCAAGAGATGGTAC 609
181 ArgValThrSerAspGlyMetLeuThrLysLysLeuIleGluArgMetValArgThrAsp 200
610 CGAGTCACCTCTGATGGCATGCTGTGGAAGAAGCTTATCCAGAGAATGGTCAGGACAGAT 669
201 SerLeuThrArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrgLysAsnLys 220
670 TCTCTGTGGAGAGCGCTGCGAAGCAAGAGATGGGACAGTATTTATTCAAAACAA 729
221 ProProAspGlyAsnAlaProProAsnSerPheTyrgAlaLeuTyrgProLysIleIle 240
730 CTTCTGTGACGGGAATGCTCTCCCACTCTTTTATAGAGCACTTTATCCTAATAATTATA 789
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
790 CAAGACATTGAGACAAATAGAACTTAATGGAGATGTGGAAGACATAGTTTACAGAGAATT 849
261 HisCysArgSerGluThrSerLysGlyValTyrgCysLeuGlnTyrgAspGlnLysIle 280
850 CACTGCGCAAGTGAACCAAGAGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA 909
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
910 GTAAGCGGCTTCGAGCAACACAAATCAGATCTCGGATTAATAACACATTGGAATGCCAG 969

QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrgAspGluArgValIle 320
Db 970 CGAATTCACAGGCCATACAGTTTCAGTCTCTGTCTCCAGTATGATGAGAGAGTATC 1029
QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 1030 ATACAGAGATCATCGGATTCACGGTCAGAGTGTGGGATGTAATAACACAGGTGAATGCTA 1089
QY 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db 1090 AACACGTTGATTACCATTTGTGAACAGTTCTGCACCTTCGCTTTCAATTAATGGCATGATG 1149
QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
Db 1150 GTGACCTGCTCCAAGATCGTTCATTCGTGTATGGGATATGGCTCCCACTGACATT 1209
QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
Db 1210 ACCCTCCGGAGGGTGTGTGTCGACACCGAGCTGTCTCAATGTGTAGACTTTGATGATGAC 1269
QY 401 LysTyrgIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
Db 1270 AAGTACATTTCTTCGATCTCGGATAGAACTATAAGGTATGGACACCAAGTACTTGT 1329
QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrgArgAsp 440
Db 1330 GAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTCAGTACAGGAC 1389
QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
Db 1390 AGCTGTGTAGTGTGAGTGGCTCATCTCACCAACTATCAGATTATGGACATAGAAATGTGT 1449
QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
Db 1450 GCATGTTTACAGTGTAGAGGCCATGAGGAATTTGTGCGTGTATTTCGATTGATTAAC 1509
QY 481 LysArgIleValSerGlyAlaTyrgAspGlyLysIleLysValTrpAspLeuValAlaAla 500
Db 1510 AAGAGGATAGTCAGTGGGCTTATGATGAAAAATAAAGTGTGGGATCTTGTGCTGCT 1569
QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
Db 1570 TTGACCCCCGTGCTCTCGAGGACACTCTCTACGACCCCTTGTGGAGCATTCGGA 1629
QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
Db 1630 AGAGTTTTTCGACTACAGATTTGATGAATTCAGATTGTCAGTAGTTCCATCATGACACA 1689
QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
Db 1690 ATCTCATCTGGGACTTCTTAATGATCCAGCTGCCAAGCTGAACCCCGGTTCCCT 1749
QY 561 SerArgThrTyrgThrTyrgIleSerArg 569
Db 1750 TCTCGAACATACACTACATCTCCAGA 1776

RESULT 6

AF101784 1818 bp mRNA linear PRI 21-JAN-1999
LOCUS Homo sapiens b-TRCP variant E3RS-IkappaB mRNA, partial cds.
DEFINITION AF101784
ACCESSION AF101784.1 GI:4165135
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1818)
Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A. M.,
Andersen, J. S., Mann, M., Mercurio, F. and Ben-Neriah, Y.
Identification of the receptor component of the
IkappaBalpha-ubiquitin ligase

```

AL Nature 396 (6711), 590-594 (1998)
NE 99075339
ED 9859996
CE 2 (bases 1 to 1818)
RS Yaron, A., Hatzubai, A., Mercurio, F., Mannig, A. M., Andersen, J. S.,
Mann, M. and Ben-Neriah, Y.
AL Direct Submission
Submitted (27-OCT-1998) Immunology, Hebrew University-Hadassah
Medical School, Jerusalem, Israel
S Location/Qualifiers
   1..1818
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   <1..1818
   /note="beta-transducin repeats-containing protein;
   receptor component of IkappaBa-ubiquitin ligase; contains
   F-box and WD regions"
   /codon_start=1
   /product="b-TRCP variant E3RS-IkappaB"
   /protein_id="AAD08702.1"
   /db_xref="GI:4165136"
   /translation="MDPADVLQERALKFMCSPRSLWLGSSLDMSPLRCLYNPG
   TGAATAFONSSREDCNCEPPRKIIPEKNSLROTYNSCARLCLNQETVCLASTAMKT
   ENCVAKTLANGTSSMIVPKORKLASVEKEKELCVKFEOMSESDOFEVFEHLISOM
   CHYQGHINSYKPKLQDFIATLARGLDHIAENILSYLDKSLCAELVCKEYRV
   TSDGMLKKLIERWTDLSGLAERGGWGYLFKNKPPDGNAPNSFYALYFKII
   QDIETESNRCGRHSIORHCRSTSGVGYCLYDQKIVSGLRDNTIKWIKNTLE
   CKRIILHTGTVLCIQYDERVITIGSSDSVAVMDVNTGEMTLIHRCEAVLHLRFN
   NGMWYTCSDRSIAVMDASPTDITLRLVLGHRAAVNVDPDDKYIVSAGSDRTIKV
   WNTSCFVRTLNGHKGKACIQYDRDLVWSSDNTIRLNDIECGACLRVLEGHEEL
   VRCIPDNKRIYVSGAYDCKIKWDLVAALDPPAPAGTICLTLVHSESRVRLQDFEF
   QIVSSSHDDTIILWFLNDPAAQAPRSPSIYTIYSR"
UNT 532 a 385 c 445 g 456 t

nt Scores:
O.: 1-25e-271 Length: 1818
Similarity: 3006.00 Matches: 569
cal Similarity: 94.05% Conservative: 0
attach: 94.05% Mismatches: 0
          99.08% Indels: 36
          9 Gaps: 1

01-168B-2 (1-569) x AF101784 (1-1818)
1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
1 ATGGACCCGGCGAGCGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGTGCTTATGCC 60
16 ----- 16
61 AGGTCTCTGTGGCTGGGCTGCTCCAGCTGGCGGACAGCATGCTTCGCTGGATGCGCTG 120
17 -----AsnSerSer:GluArgGluAspCys 24
121 TATAATCCAGGACTGGCGCACTCAGAGCTTTCCAGAAATTCCTCAGAGAGAGAACTGT 180
25 AsnAsnGlyGluProArgLysLlelleProGluLysAsnSerLeuArgGlnThrTyr 44
181 AATAATGGCGCAACCCCTAGGAAGATAATACAGAGAAGAAATTCACCTTAGACAGATAC 240
45 AsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMet 64
241 AACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTAGTATGTTAGCAAGCACTGCTATG 300
65 LysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVal 84
301 AAGACTGAGAAATTGTGTGGCCAAAACAAACCTTGCCAAATGGCACTTCCAGTAGATGTG 360
85 ProLysGluArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCysValLysTyr 104
361 CCCAAAGCAACGGAAACTCTCAGCAAGCTATGAAAGGAAAGGAAAGGAACTGTGTGTCAAATAC 420

```

```

QY 105 PheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSerGlnMet 124
DB 421 TTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGATTTGTGGAAATCTTATATCCCAATG 480
QY 125 CysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPhe 144
DB 481 TGTCAATTACCACCATGGGCACATAAACTCGTATCTTAAACCTATGTTGCAGAGAGATTTC 540
QY 145 IleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeu 164
DB 541 ATTAAGTCTCTGCCAGCTCGGGATTTGGATCATATCGCTGAGAACATTTCTGTACACCTG 600
QY 165 AspAlaLysSerLeuCysAlaAlaGluValCysLysGlnTrpTyrArgValThrSer 184
DB 601 GATGCCAAATCACTATGTGCTGCTGAACCTTGTGTCAAGGAATGGTACCGAGTGACCTCT 660
QY 185 AspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpArg 204
DB 661 GATGGCATGCTGTGGAGAAGCTTATCGAGAGAAATGGTCAGGACAGATTTCTGTGGAGA 720
QY 205 GlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspGly 224
DB 721 GGCTGGCACAACGAAGAGGATGGGACAGTATTTATTCAAAAACAACCTCTCTGACGG 780
QY 225 AsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIlelleGlnAspIleGlu 244
DB 781 AATGCTCTCTCCCAACTCTTTTATAGAGCACTTTATCTCTAAATTTATACAAGACATTGAG 840
QY 245 ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
DB 841 ACAATAGAAATCTAATTTGGAGATGTGGAAGACATAGTTTACAGAGAAATTCAGTCCGAA 900
QY 265 GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu 284
DB 901 GAAACAAGCAAGAGGATTTACTGTTTACAGATGATGATGATGATGATGATGATGATGATGAT 960
QY 285 ArgAspAsnThrIleLysLysLysLeuTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
DB 961 CGAGACAACAACAATCAAGATCTGGGATAAATAAACAACAATGGAAATGCAAGCGAAATCTCACA 1020
QY 305 GlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIlelleThrGlySer 324
DB 1021 GGCCATACAGTTTCAGTCTCTGCTCCAGTATGATGAGAGAGTGAATCAATACAGGATCA 1080
QY 325 SerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIle 344
DB 1081 TCGGAATTCACGGTTCAGAGTGGGATGTAATATACAGGTCAAAATGCTAAACACGTTGATT 1140
QY 345 HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer 364
DB 1141 CACCATTTGAAGCAGTTTCGACCTTGCCTTTCAATAATGGCATGATGGTACCTGCTCTCC 1200
QY 365 LysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArg 384
DB 1201 AAGATCGTTCATTCCTGATGGGATATGGCTCCCCAACTGACATTTACCTCCGGAGG 1260
QY 385 ValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleVal 404
DB 1261 GTGCTGTGGCAGACCGAGCTGCTGTCATGTTGTAGACTTTGATGACAAGTACATGTTT 1320
QY 405 SerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArg 424
DB 1321 TCTGCAATCTGGGATAGAACTATAAAGGTATGAAACACAAGTACTTTGTGAATTTGTAAAG 1380
QY 425 ThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVal 444
DB 1381 ACCTTAATGGACACAAACGAGGCATTTGCTGTTTCAGTACAGGACAGGCTGGTAGTG 1440
QY 445 SerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArg 464
DB 1441 AGTGGCTCTCTGACAACACATATCAGATTTATGGACATAGAAATGTGTGTCATGTTTACGA 1500

```

```

465 ValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleVal 484
1501 GTGTTAGAGGCCATGAGGAATGTGTGCTTGTATCGATTGATTAACAGAGGATATC 1560
485 SerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaLeuAspProArg 504
1561 AGTGGGGCCATGATGGAATAATTAAGTGTGGATCTGTGGCTGCTTTGACCCCGT 1620
505 AlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArg 524
1621 GTCCTCGAGGACACTCTGTCTACGACCCCTGTGGAGCATTCGGAGAGTATTTTCGA 1680
525 LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrp 544
1681 CTACAGTTTGATGAATTCAGATTGTCTAGTAGTTCACATGATGACACATCTCATCTGG 1740
545 AspPheLeuAsnAspProAlaGlnAlaGlnAlaGluProProArgSerProSerArgThrTyr 564
1741 GACTTCCATAATGATCCAGCTGCCAAGCTGAACCCCGTTCCTCCCTTCTCGAACATAC 1800
565 ThrTyrIleSerArg 569
1801 ACTACATCTCCAGA 1815

```

```

BC027994 2308 bp mRNA linear PRI 01-MAY-2002
Homo sapiens, beta-transducin repeat containing, clone MGC:40028
IMAGE:5180993, mRNA, complete cds.
BC027994
BC027994.1 GI:20380815
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2308)
Direct Submission
Submitted (08-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@hgr.nih.gov
Ahter,N., Ayala,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouford,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granate,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Laric,P., Legaspi,R., Maduro,Q.L.,
Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
Pearson,R., Stantripp,S., Thomas,P.J., Touchman,J.W., Tsugeon,C.,
Vogt,J.L., Walker,M.A., Wecherby,K.D., Wiggins,L., Young,A.,
Zhang,L.-H. and Green,E.D.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 62 Row: e Column: 7

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502476.

Location/Qualifiers

```

1..2308
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:8945"

```

```

/db_xref="taxon:9606"
/clone="MGC:40028 IMAGE:5180993"
/tissue_type="Brain, lung, testis, adult, pooled whole"
/clone_lib="NIH MGC 115"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
38..1855
/codon_start=1
/product="beta-transducin repeat containing"
/protein_id="AAH27994.1"
/db_xref="GI:20380816"
/translation="MDPAVQLQEKALFMCSPRSLWLGSSSLADSMPSRLCLNPG
TGAITAFONSEREDCNNGEPKLIPEKNSLRQTVNSCARCLNLOETVCLASTAMKT
ENCVAKLIANGTSSMIVPQKRLSASVEKEKELCVKFEQMSDOVEFVSHLLSQM
CHYQGHINLYLKMPLORDPITALPARGLDHIAENILSYLDAKSLCAELVCKEWIRV
TSDMLWKKLIERNVRLDSLWGLAERRGGYLFKNKPPDGNAPNSFYRLYPKII
QDIETESNWRGSRHSLOIRHCRSETSGVYLOYDDQKIVSLGRDNTIKIWDKNTLE
KRIITGSGVLCIYDRIITGSDSTVWVDVNTGEMLNTLIHHCBAVLHILFN
NGMWTCSFVRIINGHKGRIACLYDRILNVSGSSDNTIRLMDIEGACLRVLEGEHEL
WNTSTCFVRTINGHKGRIACLYDRILNVSGSSDNTIRLMDIEGACLRVLEGEHEL
VRCIRFENKRIYSGAYDGKIKYWDIYVAALDPRAPAGTICLRILVEHSGRVFRLQDFE
QIVSSSHDDTILWDFLNDPAQAEPSPSKTYITISR"
BASE COUNT 719 a 493 c 530 g 566 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1.73e-271 Length: 2308
Score: 3006.00 Matches: 569
Percent Similarity: 94.05% Conservative: 0
Best Local Similarity: 94.05% Mismatches: 0
Query Match: 99.08% Indels: 36
DB: 9 Gaps: 1

US-09-601-168b-2 (1-569) x BC027994 (1-2308)
Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
Db 38 ATGGACCGCGCGAGCGGTGCTGCAAGAGAGGCACTCAAGTTATGTGCTCTATGCC 97
Qy 16 ----- 16
Db 98 AGGTCTCTGTGGCTGGGCTGCTCAGCTGGCGGACAGCATGCTTGGCTGGATGCTG 157
Qy 17 ----- AsnSerSerGluArgGluAspCys 24
Db 158 TATAACCCAGGACTGGCGCACTCACAGCTTTCAGAAATCTCAGAGAGAGAACTGT 217
Qy 25 AsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerIleuArgGlnThrTyr 44
Db 218 AATAATGGCGAACCCCTAGGAAGATAATACCAGAGAGAAATTCACCTTAGACAGACATAC 277
Qy 45 AsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysIleuAlaSerThrAlaMet 64
Db 278 AACAGCTGTGCCAGACTCTGTTAAACCAAGAAACAGTAGTATTAGCAAGCACTGCTATG 337
Qy 65 LysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVal 84
Db 338 AAGACTGAGAATTGTGTGGCCAAACAAACCTTGCATGGCACTTCCAGTATGATTGTG 397
Qy 85 ProLysGluArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCysValLysTyr 104
Db 398 CCCAAGCAACGGAAACTCTCAGCAGCTATGAAAGAGAAAGAACTGTGTGTCAAAATAC 457
Qy 105 PheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisIleuIleSerGlnMet 124
Db 458 TTGACAGTGTGTGAGTCAGATCAAGTGGAAATTTGTGGAACATCTTATATCCCAATG 517
Qy 125 CysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPhe 144
Db 518 TGTCAATACCAATGGGCACATAACTCGTATCTTAAACCTATGTTGAGAGAGATTTC 577
Qy 145 IleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeu 164

```

578 ATRACTGCTGCGCAGCTCGGGGATTGGATCATATTGCTGAGAAACATTCTGTACATCCTG 637
 165 AsphAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSer 184
 638 GATCCCAAAATCACTATGCTGCTGCTGAACCTTGCTGCAAGGAATGGTACCGAGTACCTCT 697
 185 AspGlyMetLeuTrpLysLysLeuLeuGluArgMetValArgThrAspSerLeuTrpArg 204
 698 GATGGCATGCTGTGGAGAGAGCTTATCGAGAGAAATGGTCAGGACAGATTCTCTGTGGAGA 757
 205 GlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGly 224
 758 GGCCTGGCAGAACGAGAGGATGGGACAGTATTATTCAAAACAAACCTCTGACGGG 817
 225 AsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysLeuLeuGlnAspLeuGlu 244
 818 AATGCTCTCCCACTCTTTTATAGAGCACTTTATCCPAAATATATACAGACATTGAG 877
 245 ThrLeuSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
 878 ACAATAGAACTAATTGGAGATGGGAAGACATAGTTTACAGAGAATTCACCTGCCGAAGT 937
 265 GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysLeuValSerGlyLeu 284
 938 GAAACAGCAAGAGGTTTACTGTTTACAGTATGATGATCAGAAATATAGTAGCGCCTT 997
 285 ArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
 998 CGAGACAACAACATCAAGATCTGGGATATAAAACACATTTGGAATGCAAGCGAATTTCTACA 1057
 305 GlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySer 324
 1058 GGCATACAGGTCAGTCCTCTGTCTCCAGATGATGATGAGAGAGTATATACAGAGTCA 1117
 325 SerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIle 344
 1118 TCGGATTTCCACGGTCAGAGTGGGATGTAATACAGGTGAATGCTAAACACGTTGATT 1177
 345 HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer 364
 1178 CACCATTTGTAAGCAGTTCTGCACTTTCGCTTCAATATGGCATGATGTCACCTGCTCC 1237
 365 LysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArg 384
 1238 AAGATCGTTCCATTGCTGTATGGGATATGCGCTCCCACTGACATTTACCTCCGGAGG 1297
 385 ValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleVal 404
 1298 GTGCTGTGCGACACCGAGCTGCTGCTCAATGTTGTAGACTTTGATGACAAGTACATTGTT 1357
 405 SerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArg 424
 1358 TCTGCATCTGGGATAGAACTATTAAGTATGGAACACAAAGTACTTGTGAATTTGTAAGG 1417
 425 ThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVal 444
 1418 ACCTTAATGACACAAACAGGAGCATGCTGTTTTCAGTACAGGACAGGCTGGTAGTG 1477
 445 SerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArg 464
 1478 AGTGGCTCATCTGACACACTATCAGATTATGGGACATAGAAATGTTGGTGCATGTTTACGA 1537
 465 ValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleVal 484
 1538 GTGTTAGAGGCCATGAGGAATTTGGTGGTGTATTCGATTGATTAACAGAGGATAGTCT 1597
 485 SerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArg 504
 1598 AGTGGGCGCTATGATGAAAAAATTAAGTGTGGGATCTTGTGCTGCTTTGGACCCCGCT 1657
 505 AlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArg 524
 1658 GCTCTTGCAGGGACACTCTGTCTACGGACCCCTTGTGGAGCAATTCCTCGGAAGAGTTTTTGA 1717

Qy 525 LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrp 544
 Db 1718 CTACAGTTTGATGAATTCAGAGATTGTAGTAGTTTCACATGATGACACAAATCTCATCTGG 1777
 Qy 545 AspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerProSerArgThrTyr 564
 Db 1778 GACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCGTTCCCTCTCTCGAACATAC 1837
 Qy 565 ThrTyrIleSerArg 569
 Db 1838 ACCTACATCTCCAGA 1852
 RESULT 8
 AF099932 1712 bp mRNA linear ROD 21-JAN-1999
 LOCUS Mus musculus beta-TrCP protein E3RS-IkappaB mRNA, complete cds.
 DEFINITION AF099932
 ACCESSION AF099932
 VERSION AF099932.1 GI:4008019
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1712)
 Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A.M.,
 Andersen, J.S., Mann, M., Mercurio, F. and Ben-Neriah, Y.
 Identification of the receptor component of the
 IkappaBalpha-ubiquitin ligase
 Nature 396 (6711), 590-594 (1998)
 JOURNAL 98075339
 MEDLINE 9859996
 PUBMED
 REFERENCE 2 (bases 1 to 1712)
 Yaron, A., Hatzubai, A., Mercurio, F., Manning, A.M., Andersen, J.S.,
 Mann, M. and Ben-Neriah, Y.
 Direct Submission
 Submitted (19-OCT-1998) Immunology, Hebrew University of Jerusalem,
 Ein Karem, Jerusalem 91120, Israel
 FEATURES
 Location/Qualifiers
 1..1712
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1..1710
 /notes="beta-transducin repeat-containing protein; receptor
 component of IkBa-ubiquitin ligase"
 /codon_start=1
 /product="beta-TrCP protein E3RS-IkappaB"
 /protein_id="AAD08701.1"
 /db_xref="GI:4008020"
 /translation="MDPAEAVLOEKALFMNNSREDCNNGEPPPKIIPKNSLRQTY
 NSCARLINCQETVCLTSTAMKTCNCVAKAKLANGTSSMIVPKOKLSASYEKELCV
 KYFEQMSDQVEFVHEHLSIQMCHYQGHINSLYKAPMLQRFITALPARGDQHAENI
 LSYLSDASQAEALVCKEWRYVSDGMLKHLIERMVRTSLMRGLAERRGQGLYFK
 NKPPDNAPNSFYALYPKITQDIETTESNRCGRHSLOIHCRESYSGVYCLQYD
 DQKIVSGLDNTIKWDSLTKEKLTGTCVLCQYDERVITGSSDSTVRVVDV
 NAGEMLTUHVCEASVHLRFNNGMVTCSKDRSAVWMASTPTILRRVLVGHRAA
 VNVVDKDKVIVSASGDRITKVNTSTCFVRTLNGHKGAGIACLOYRDLVVGSSDN
 TIRLWDIEGACGLRVLEGHEELVRCIRDNKIVSGAYDGGKIKWMDLMAALDPRAPAG
 TICRLTLVSHSRVFRQLQDFEQIVSSSHDDTLILWDFNDPAAHAEPSPSRITY
 ISR"
 BASE COUNT 469 a 399 c 453 g 391 t
 ORIGIN
 Alignment Scores: Length: 1712
 Pred. No.: 8e-271 Matches: 561
 Score: 2997.00
 Percent Similarity: 99.12% Conservative: 3
 Best Local Similarity: 98.59% Mismatches: 5
 Query Match: 98.78% Indels: 0
 DB: 10 Gaps: 0
 US-09-601-168B-2 (1-569) x AF099932 (1-1712)

```

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
1 ATGGACCGCGCGAGCGGTGTCGAGGAGAAGCGCTTAAGTTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
61 AGAGAAGACTGTAATAATGGCGAACCCCTAGGAAGATAATACNAGAGAAGAAATTCATT 120
41 ArgGlnThrThrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACTTACAACAGCTGTGCCAGGCTTTGCATAAACCAGAGACAGATATGCTAACA 180
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGCTATGAAGACTGAATAATGTGTGGCCAAAGCCAACTTGGCCAAATGGCACTTCC 240
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrTyrGluLysGluLysGluLeu 100
241 AGCATGATTGTGCCCAAGCAGCGGAACCTCTCAGCAAGCTATGAGAAGAAAGGAGCTG 300
101 CysValLysTyrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
301 TGTGTCAAGTATTTTTCAGCAGTGGTCAGAGTCTGATCAAGTGGAAATTTGTAGAACACCTT 360
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
361 ATATCCCAATGTGTCACTACCAAGCTGGGACATCAACTCTACTTAACCTATAGCTG 420
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
421 CAGAGGATTTTCATNACTCACTGCCAGCAGCAGCGGTCTGGACCATCGCTGAGAACATT 480
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluThrPyr 180
481 CTGTCAATCTTGGACGCCAAGTCACTGTGTGTCTGTAGCTGTGTGCAAGAAATGGTAC 540
181 ArgValThrSerAspGlyMetLeuThrLysLysLeuIleGluArgMetValArgThrAsp 200
541 CGCGTCACGTCGGACCGCATGCTGTGGAAAAGCTCATCGAGAGGATGGTCAGGACGGAC 600
201 SerLeuThrArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
601 TCTCTGTGGCGAGCGCTGGCAGAGCGCAGAGGCTGGGGACAGTACTTATTAAAAACAAA 660
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
661 CTCTCTGATGAGAACGCTCTCCCACTCTCTTTATAGACGCTTTATCCTTAATAATCAT 720
241 GlnAspIleGluThrIleGluSerAsnThrArgCysGlyArgHisSerLeuGlnArgIle 260
721 CAAGACATTGAGACAATAGAGTCCAATTTGGAGATGTGGCGACATAGTTTACAGAGAATC 780
261 HisCysArgSerGlnThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
781 CACTGCGGAGTGAACAAGTAAAGGGGTTTACTGTTTACAGTACGACACAGAGAGATA 840
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
841 GTCAGGGGCTTCGAGACACACCATCAAGATCTGGGATAAAGACACACTGGAATGCCAG 900
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
901 CGGATTCTCAGCGGCCACACGGGCTCCGCTCTGTCTGCTGCTGATGAGAGGGGTGATC 960
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
961 ATCAGAGGCTCCCTCAGACTCCACCGTCAGAGTGTGGGATGAATTCAGGTGAGATGCTA 1020
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1021 AACACATTGATTCAACCACTGTGAAGCGGCTTCTGCACCTGGGCTTCAATATGGCATGAT 1080
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1081 GTGACCTGTTCCAAAGACCGTTCCATCGTGTGGGATATGGCTTCCCACTGACATC 1140
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1141 ACCCTCAGGAGGGTGTGTGTGGGACACCGAGCTGCGGTCAATGTTGTAGACTTTTGATGAC 1200
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1201 AAGTACATCGTTTCTCGCTCTGAGATAGAACAATAAAGGTGTGGAACACAAGTACCTGT 1260
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1261 GAATTGTTAAGACCTTAANTGGGCACAAGCGTGCATCGCTGTTTGCAGTACAGAGAC 1320
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1321 AGGCTGGTGTGAGCGGCTCTCTGACAACACCATCAGCTGTGGGACATAGAGTGTGGA 1380
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1381 GCATGCTTCGAGTGTGTGGAGGGCCATGAGGAGTTGGTACGCTGCATTCGATTTCATAAC 1440
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
1441 AAAAGATAGTGAGCGGAGCCCTATCATGGGAAAATTAAGTGTGGGATCTTATGCTGCT 1500
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1501 TTGACCCCGCTGCTCCAGCAGGACTCTCTGCTCTGCGGACACTTGTGGAGCAATCTCGA 1560
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540
1561 AGATTTTTCGCTCCAGTTTGTGAATTCAGATTTGCAGTAGTTCATATGATGACACA 1620
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1621 ATTCTCATCTGGAGCTTCTGTAATGATCCAGCTGTCTCAGCTGAACCGCCGCTCCCT 1680
561 SerArgThrTyrThrTyrIleSerArg 569
1681 TCTCGACATACACCTTACATCTCCAGA 1707

```

BC003989 2081 bp mRNA linear ROD 16-APR-2003
Mus musculus beta-transducin repeat containing protein, mRNA (cDNA
Clone MGC:7517 IMAGE:3491843), complete cds.

BC003989
VERSION BC003989.1 GI:13278339
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

RESULT 9
BC003989
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 2081)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Shpetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Mavris, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 2081)
 CE Strausberg, R.
 RS Direct Submission
 AL Submitted (28-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 K NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb-remail.nih.gov
 Tissue procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 8 Row: 0 Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6753209.
 Location/Qualifiers

S

```

1..2081
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="XREF:7517 IMAGE:3491843"
/tissue_type="mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MMTV."
/clone_lib="NCI CGAP_Lu29"
/lab_host="DH10B"
/vector="PCMV-SPORT6"
1..2081
/gene="Btrc"
/note="synonyms: FWD1, Slimb, b-TrCP, mSlimb"
/db_xref="LocusID:12234"
/db_xref="GI:13278340"
7..1716
/codon_start=1
/product="beta-transducin repeat containing protein"
/protein_id="AAH03989.1"
/db_xref="GI:13278340"
/db_xref="LocusID:12234"
/translation="MDPAEVLQEKALPMNSSEDCNNGEPKRIPEKNSLRQV
NSCARLCTNOETVCLTSTAMKTCVAKLANTSGTIVPQKRLSAYEKELCV
KFEQWSDQVEFHEILSQMCHQHQHGLYILKPMLOQDFITAPALPARGLDHIAENI
LSYLAQSKAELAEVCKWRYVTSQGLWKKLIETRWETSLMRGLAERRGWQYLFK
NKPPENAPPNSFYALPKYKIIQDIETIESNRCRHSQRIHCRSETSKGVCLQYD
DQKIVSGRLNDIKWDSLTLECKRIILGHTGSLVLCQYDERVITGSSDSTVRWDV
NAGEMLNTLIHCEAVLHLRFNNGMWVTCSDKRSIAVWDMASPTDITLRRVLVGRRAA
VNVDPFDKVIIVASGSDRTIKVYNTSTCFEVRTLNHKGRIACIQYDRDLVSSSDN
TIRLWDIEGACLRVLEHGEELVRCIRDNKRIVSGYDQKIKVWDMALDPPAPAG
TICLRILVHSGSRFRLQDFEQIVSSSHDDTILIDFLNDPAAHAPSPSPSRTY
ISR"
JNT 568 a 502 c 529 g 482 t

```

at Scores:

1.05e-270 Length: 2081
 2997.00 Matches: 561

```

Percent Similarity: 99.12% Conservative: 3
Best Local Similarity: 98.59% Mismatches: 5
Query Match: 98.78% Indels: 0
DB: 10 Gaps: 0
US-09-601-168B-2 (1-569) x BC003989 (1-2081)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 7 ATGGACCCGCGCAGAGCGGTGCTGCAGGAGAAAGCGCTTAAGTTTATGAATTCCTCAGAG 66
Qy 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
Db 67 AGAAGAAGACTGTAATAATGCGAAACCCCTAGGAGATTAATACCCAGAGAAGAAATTCAC 126
Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 127 AGACAGACTTACACAGCTGTGCCAGCTTTCATTAACCAAGACAGACAGTATGCTTAACA 186
Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 187 AGCAGCTGCTATGAAGACTGAAATTTGTGTGGCCAAAGCCAAACTTGCCCAATGSCACTCC 246
Qy 81 SerMetIleValProLysGlnArgLysIleSerAlaSerTyrGluLysGluLysGluLeu 100
Db 247 AGCATGATGTGTCGCCAGCAGCGGAACTCTCAGCAAGCTATGAGAGGAAAGAGAGCTG 306
Qy 101 CysValLysTyrPheGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 307 TGTGTCAAGTATTTTTCAGCAGTGTGTCAGAGTCTGATCAAGTGGAAATTTGTAGAACCTT 366
Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
Db 367 ATATCCCAATGTGCTACCTACCGCATGGGCACTCACTCTCTATAAACTATGCTG 426
Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
Db 427 CAGAGGATTTTCACTACTGCATCCAGCAGCAGCGGTCTGGACCACTATCGCTGAGAACAT 486
Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr 180
Db 487 CTGTCATCTTGGACGCCAAGTCACGTGTGCTGTGCTGAGCTCGTGTGCAAGGAATGGTAC 546
Qy 181 ArgValThrSerAspGlyMetLeuThrLysLysLeuIleGluArgMetValArgThrAsp 200
Db 547 CGCGTACGTCGCGACCGCATGCTGTGAAAAAAGCTCATCGAGAGGATGTCAGGACGGAC 606
Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
Db 607 TCTCTGTGGCGAGGCTTGGCAGAGCGCAGAGGCTGGGGACAGTACTTATCAAAAACAAA 666
Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
Db 667 CCTCCTGATGAGAACGCTCTCTCCAACTCTTTTATAGAGCGCTTTATCCTAAAAATCATA 726
Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db 727 CAAGACATTGAGACAAATAGAGTCCAAATGAGATGTCGGGACATAGTTTACAGAAATC 786
Qy 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
Db 787 CACTGCCGAGTGAACAAGTAAGGGGTTTACTGTTTACAGTACGACGACGACAGAGATA 846
Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
Db 847 GTCAGGGCGCTTCGAGACAAACACCATCAAGATCTGGGATAAAGACACACTGGAATGCAAG 906
Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 907 CGGATTCTCAGGGGCCACACGGGCTCCGTCCTGTGTCAGTACGATGAGAGGGGTGATC 966
Qy 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340

```


967 ATCAGAGGCTCCTCAGACTCCACCGTCAGAGTGGGATGTAATGCGAGGTGAGATGCTA 1026
341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1027 AACACATTGATTCCACCACTGTGAAGCGGTTCTGCACTGCGCTTCAATAATGGCATGATG 1086
361 ValThrCysSerLysArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1087 GTGACCTGCTCCAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCAACTGACATC 1146
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1147 ACCCTCAGAGGGTCTGTGTGGGACCGGACCGGCTGGGTCAGTGTGTAGATTTGTAGTAC 1206
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1207 AAGTACATCGTTTCTGCTCTGGAGATAGACCAATAAGGTGTGGACACACAGTACCTGT 1266
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1267 GAATTCTGAAGACCGCTTAATGGGCACCAAGCGTGGCATCGGCTGTTTGCAGTACAGAGAC 1326
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1327 AGCGTGTGTGGAGCGGCTCTCTGACACACCATCAGGCTGTGGACATAGATGTGGA 1386
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1387 GCATGCTCGGAGTGTGGAGGGCCATCAGGAGTGTGTACGCTGCATTCGATTGTGATAAC 1446
481 LysArgIleValSerGlyValAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
1447 AAAAGGATAGTGGAGCGGAGCTATGATCGGAAATAAAGTGTGGGATCTTATGCGTCT 1506
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1507 TTGGACCTCGTCTCCAGCAGGAGTCTCTGTCTGGGACACTGTGGGACATCTGGA 1566
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
1567 AGAGTTCCTCGCTCCAGTTTATGATGATTCAGATGTCAGTGTTCACATGATGACACA 1626
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1627 ATTCTCATCTGGAGCTTCTGAATGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
561 SerArgThrTyrThrTyrIleSerArg 569
1687 TCTCGGACATACCTACATCTCCAGA 1713
0
AF081887 2175 bp mRNA linear ROD 07-OCT-1999
ON Mus musculus ubiquitin ligase FWD1 mRNA, complete cds.
DN AF081887
AF081887.1 GI:4336326
3 Mus musculus (house mouse)
SM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2175)
Hatakeyama, S., Kitagawa, M., Nakayama, K., Shirane, M., Matsumoto, M.,
Hattori, K., Higashi, H., Nakano, H., Okumura, K., Onoe, K., Good, R. A.
and Nakayama, K.-I.
Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
ubiquitin ligase Skp1/Cul 1/F-box protein FWD1
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3859-3863 (1999)
99199275
JE 10097128
DE 2 (bases 1 to 2175)
E Hatakeyama, S. and Nakayama, K.-I.
S Direct Submission

JOURNAL Submitted (04-AUG-1998) Department of Molecular and Cellular
Biology, Medical Institute of Bio-regulation, Kyushu University,
3-1-1, Maedashi, Higashi-Ku, Fukuoka 812-8582, Japan
FEATURES
source
1. .2175
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
91..1800
/note="similar to beta-transducin repeat-containing
protein Trcp; F-box protein; contains WD40 repeats"
/codon_start=1
/product="ubiquitin ligase FWD1"
/protein_id="AAD17755.1"
/db_xref="GI:4336327"
/translation="MDPAEAVLOEKALFMSSBERDCNNGEPKIIPEKNSLRCTY
NSCARLQNQETVCLTSTAMKTNCHVAKALANGYSSMIVPKOKLSAYEKELCV
KTFEWSMSDQVEFVHLISQHGHIKNSQVLMQORFITALPARGLDHIAENI
LSYDASLCAABLVCBMYRVTSQMLKWKLIERMVTRDSLMRGLAERGQYLFK
NKPPENAPPSFYRALYPKITQDIETIESNWRGHSQRIHCRSETSKGVYCLQYD
DOKIVSGLRDNTIKWDXSTLEKRLITGHTGSLCVQDERVITGSSDSSTVRVMDV
NAGEMLTLIHCEAVLHLRNNGMVTCSDRSIAVMDASPTDITLRRVLVGHRAA
VNVDPDKYIVYASGDRITKIVNTSTCEFTVTLNKHKGRIACLOYRDLVVGSGSDN
TIRLMDIECGACIRVLEHEELVRCIFRDNKRIVSGAYDGKIKWMDLMAALDPRAPAG
TICLRLTVEHSGRVFRLOQDFEQIVSSSHDDTILIDFLNDFAAHAEPSPSRITY
ISR"
BASE COUNT 589 a 528 c 567 g 491 t
ORIGIN
Alignment Scores:
Pred. No.: 1,116-270 Length: 2175
Score: 2997.00 Matches: 561
Percent Similarity: 99.12% Conservative: 3
Best Local Similarity: 98.59% Mismatches: 5
Query Match: 98.78% Indels: 0
DB: 10 Gaps: 0
US-09-601-168B-2 (1-569) x AF081887 (1-2175)
QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 91 ATGGACCCGGCAGAGCGGTGTGTGAGGAGAAAGCGTTAAGTTATGAAATTCCTCAGAG 150
QY 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
Db 151 AGAGAAGACTGTATAATGGCAACCCCTAGGAAGATAATACCAAGAGAAATTCACCT 210
QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 211 AGACAGACTTTACACAGCTGTGCCAGGCTTTCATAAACCAAGACACAGATGTATGCTTACA 270
QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 271 AGCAGTGTATGAGACTGAAATTTGTGTGGCCAAAGCCAAACTTGCACATGGCACTTCC 330
QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db 331 AGCATGATTGTCCCAAGCAGCGGAAACTCTCAGCAAGCTATGAGAAGAAAGGAGCTG 390
QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 391 TGTGTCAAGTATTTTGTAGCAGTGTGTCAGAGTCTGATCAAGTGAATTTGTAGAACACCT 450
QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
Db 451 ATATCCCAATGTGTCACTACCAAGATGGGCACATCACTCTACTATAAACCTATGCTG 510
QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
Db 511 CAGAGGATTTTCATAACTGCATGCCAGCAGCGGGTCTGGACACACATCGCTGAGAACATT 570
QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180


```
1  ATGGACCCGGCAGAGCGGTGCTGCAGGAGAAAGCGCTTAAGTTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
61 AGAGAAGAGCTGTAAATAATGGCGAACCCCTAGGAAGATAATACCAGAGAAATTCACCT 120
41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACTTACACAGCTGTGCCAGGCTTTGCATAAACCAAGAGACAGATATGCTAACA 180
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGCTATCAGAGACTCAAAATTGTGTGSCCAAGCCAAATGTGCCAATGGCACTTCC 240
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
241 AGCATGATTGTGCCCAAGCAGCGGAACTCTCAGCAAGCTATGAGAAGAAAGAGAGCTG 300
101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
301 TGTGTCAAGTATTTTGAGCAGTGTGCAGAGTCTGATCAAGTGGAAATTTGTAGAACACCTT 360
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPheMetLeu 140
361 ATATCCCAAAATGTCTACTACAGCATGGGCACATCAACTCTCTACCTAAACCTATGCTG 420
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
421 CAGAGGGATTTCTAACTGCACTGCCAGACGCGGTCTGGACCACTGCTGAGAACTT 480
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
481 CTGTCTACTTGGACGCGCAAGTCACTGTGTGCTGCTGAGTCTGTGTCAGGAATGGTAC 540
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
541 CGCGTGACGTGGACGGGATGCTGTGGAAAAAGTCTATCGAGAGAGATGCTCAGACGGAC 600
201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
601 TCTCTGTGGCGAGGCTGGCAGAGCGCAGAGCTGGGACAACTACTTATTCAAAACAAA 660
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
661 CCTCTGTATGAGAACGCTCTCCCACTCTTTATAGAGCGCTTATCTCTAAAATCATA 720
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
721 CAAGACATTGAGACAAATAGAGTCCAAATTTGGAGATGTGGCGACATAGTTTACAGAGAATC 780
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
781 CACTCCCGAGTGAACCAAGTAAGGGGTTACTGTTTACAGTACGACGACACAGAGATA 840
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
841 GTCAAGCGGCTTCGAGACACACCACTCAAGATCTGGGATAAAAGCACACTGGAATGCAAG 900
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
901 CGGATTTCTACGGGCCACACGGGCTCCGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
961 ATCAGAGGCTCTCAGACTCCACCTCAGAGTGTGGGATTAATTCAGAGGTGAGATGCTA 1020
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1021 AACACATTGATTCAACCACTGTGAAGCGGTTCTGCACCTGCGCTTCAATAATGGCATGATG 1080
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
```

1081 GTGACCTGTTCCAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCAACTGCACATC 1140

381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400

1141 ACCCTCAGAGGGTGTCTGTGGGACACCGAGTGGGTCAATGTTGTAGACTTTTGATGAC 1200

401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420

1201 AAGTACATCGTTTCTGCCTCTCGAGATAGAACCAATAAAGTGTGGACACACAAGTACCTGT 1260

421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440

1261 GAATTCGTAAAGACCCCTAAATGGGCACAAAGCGTGGCATCGCTGTTTGCAGTACAGAGAC 1320

441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460

1321 AGCTCGTGTGGAGCGCTCTCTGACAAACACCATCAGGCTGTGGACATAGAGTGTGA 1380

461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480

1381 GCATGCTCGAGTGTGGAGGCGCATGAGGAGTTGGTACGCTGCATTCGATTGATAAC 1440

481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaIle 500

1441 AAAAGGATAGTGGCGGAGCCCTATGATGGGAAAAATTTAAAGTGTGGATCTTATGGCTGCT 1500

501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520

1501 TTGGACCGCGTCTCCAGCAGGACTCTCTGTCTGGGACACTGTGGAGCATTCCTGA 1560

521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540

1561 AGAGTTTCCGCTCCAGTTTGTGATGAATTCAGATTGTCAGTAGTTCACATGATGACACA 1620

541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560

1621 ATTCTCATCTGGACTTCTGTAATGATCCAGCTGTCCAGCTGTCAAGTGAACCGCCGCTCCCT 1680

561 SerArgThrTyrThrTyrIleSerArg 569

1681 TCTCGACATACCTACATCTCCAGA 1707

RESULT 12

AF110396 1979 bp mRNA linear ROD 28-JUN-1999

LOCUS Mus musculus beta-transducin repeat-containing protein mRNA,

DEFINITION complete cds.

ACCESSION AF110396

VERSION AF110396.1 GI:5230821

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 1979)

Winston,J., Ellledge,S.J. and Harper,J.W.

Direct Submission

Submitted (01-DEC-1998) Biochemistry, Baylor College of Medicine,
One Baylor Plaza, Houston, TX 77030, USA

FEATURES

Location/Qualifiers

1..1979

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1..1710

/note="beta TRCP"

/codon_start=1

/product="beta-transducin repeat-containing protein"

/protein_id="AA041025.1"

/db_xref="GI:5230822"

/translation="MDPAEAVLQEKALPKMNSRERDCNNGEPFPRKIIPKNSLRQTY
NSCARLQINQETVCLUTSTAMTKNCVAKUANGTSSMIVPKORKLSASYEKELCV
KYFEQSSDSQVEFVHLISQMHYQGHINSYLPKMLQRFITALPARGLDHITENI

LSYLDKSLCAAELCKEWRVTSQMLWKKLIERMVRDLSMRGLAERRGWOYLFK
 NKPPDENAPNSFYALYPKIIODIETTESNRGCRHSIORIHCRSSTSGVVCLOYD
 DOKIVSGERDNTIKWDKSTLECKELTGHGTGSLVCLQYDERVLIITGSSDSTVRVMDV
 NAGEMDLTILTHCEAVLHLENFNNGMNTCSKDRSLAVHNDMSPTDILRLRVLGHRAA
 VNVDFFDDKIVISAGSDRTIKWNTSTCEFKVRLTKNKGHGLACLOYDRIVSSGSDN
 TIRLWDIECGACLRVEHELVKRIYVSGRYDGKIKYWDMLMAALDPRAPAG
 TILCLRTLVEHSGRVERFLQDFEQIVSSSHDDTILIDFLNDPAHAHPPRSPSRITY
 ISR"

UNT 535 a 480 c 516 g 448 t

nt Scores:

O.: 5.88e-269 Length: 1979
 2978.00 Matches: 557
 Similarity: 98.59% Conservative: 4
 cal Similarity: 97.89% Mismatches: 8
 atch: 98.15% Indels: 0
 10 Gaps: 0

31-168B-2 (1-569) x AF110396 (1-1979)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 1 ATGGACCCGGCAGAGCGGTGTCAGAGAGAAAGCGCTTAAGTTATGAATCTCTCAGAG 60
 21 ArgGluAspCysAsnAsnGlyGluProArgLysIlelleProGluLysAsnSerLeu 40
 61 AGAGAAGACTGTAATAATGCGCAACCCCTAGAGAGATAATACCAAGAGAAGAAATCACTT 120
 41 ArgGlnThrTyAsnSerCysAlaArgLeuLysLeuAsnGlnGluThrValCysLeuAla 60
 121 AGACAGACTTACACAGCTGTGCCAGCTTTTCATTAACCAAGACAGACTGTGTCTAACA 180
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 181 AGCACTGCTATGAAGACTGAAATATTGTGTGCCAAGCAAACTTGCCAAATGGCACTTCC 240
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyArgLysGluLysGluLeu 100
 241 AGCATGATTGTGCCCAAGCAGCGGAACCTCTCAGCAAGCTATGAGAAGGAAAGAGCGTG 300
 101 CysValLysTyPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 301 TGTGTCAAGTATTTCAGCAGTGGTCCGAGTCTGATCAAGTGGAAATTTGTAGACACCTT 360
 121 IleSerGlnMetCysHisTyrglnHisGlyHisIleAsnSerTyLeuLysProMetLeu 140
 361 ATATCCCAAATGTGTCACTACCAGCATGGGCACATCAACTCTTACCTAAACCTATGCTG 420
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 421 CAGAGGGATTTCTAACTGCACTGCCAGCAGCGGGGTCTGGACCACTCACTGAGAACATT 480
 161 LeuSerTyLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGlnTrpTy 180
 481 CTGTCACTATGGACGCCAAGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 541 CGCGTGACGTCCAGCGCATGCTGTGGAAAAAGCTCATCGAGGAGATGTCAGGACGGAC 600
 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTyTrpGlyGlnTyLeuPheLysAsnLys 220
 601 TCTCTGTGGCGAGCGCTGGCAGAGCCGAGAGCGGTGGGACAGTACTTATTCAAAACAAA 660
 221 ProProAspGlyAsnAlaProProAsnSerPheTyArgAlaLeuTyProLysIleIle 240
 661 CCTCCTGATGAGAAGCGTCTCCCAACTCCTTTTATAGACGCTTTATCCTAAAAATCATA 720
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 721 CAAGACATTGAGACATAGAGTCCAAATTGGAGATGTGGCGCACATAGTTTACAGAGAATC 780

QY 261 HisCysArgSerGluThrSerLysGlyValTyrcysLeuGlnTyrcysAspGlnLysIle 280
 DB 781 CACTGCCGAGTGAACAAAGTAAGGGGTCTACTGTTTACAGTACGAGCACCAAGAGATA 840
 QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 DB 841 GTCAGCGGCTTCGAGAGAACCAACCATCAAGATCTGGGATAAAAGCACACATGGAATGCAAG 900
 QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrcysAspGluArgValIle 320
 DB 901 CGGATTCTCAGGGGCCACACGGGCTCCGCTCTGTGTCTGCACTAGTACGAGGGGTGATC 960
 QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 DB 961 ATCAGAGGCTCTCAGACTCCACCGTCAGAGTGTGGAGTGAATATGCAAGTGAAGTGTCTA 1020
 QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisIleuArgPheAsnAsnGlyMetMet 360
 DB 1021 GACACATTACTCACCACTGTGAAGCCGTTCTGCACCTGCGCTTCAATAATGGCATGATG 1080
 QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 DB 1081 GTGACCTGTTCCAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCAACTGACATC 1140
 QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 DB 1141 ACCCTCAGAGGGTGTGTGTGGACACCGAGCTGCGGTCAATGTTGTAGACTTTTGATGAC 1200
 QY 401 LysTyrlleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 DB 1201 NAGTACATCGTTTCTGCTCTGGAGATAGAACCATTAAGGTGTGGAAACACAAAGTACCTGT 1260
 QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrcysArg 440
 DB 1261 GAAATCGTAAAGACCTTAATATGGGCACAAAGCGTGGCATCGCTGTTCAGTACAGAGAC 1320
 QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 DB 1321 AGCTGTGTGTGAGCGGCTCTCTGACACACCATCAGGCTGTGGACATAGAGTGTGGA 1380
 QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 DB 1381 GCATGCTCCGAGTGTGGAGGGCCATGAGGAGTTGGTACGCTGCATTCGATTGATAAC 1440
 QY 481 LysArgIleValSerGlyAlaTyrcysGlyLysIleLysValTrpAspLeuValAlaAla 500
 DB 1441 AAAAGGATAGTGGCGGACGCTATGATGGGAAATTAAGGTGGGATCTTATGGCTGCT 1500
 QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 DB 1501 TTGGACCGCGTGTCTCCAGCAGGACTCTCTGTCTGGGACACTTGTGGAGCATTCGGA 1560
 QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 DB 1561 AGAGTTTCCGCTCCAGTTTGTGAATTCAGATTTCAGATTGTTCAGTAGTTCACATGATGACACA 1620
 QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 DB 1621 ATTCTCATCTGGAGCTTCTCTGAATGATCCAGTGTCTCAGCTGTCTCAGCTGAACCCCGCTCCCT 1680
 QY 561 SerArgThrTyThrTyIleSerArg 569
 DB 1681 TCTCGACATACACCTACATCTCCAGA 1707

RESULT 13

XELSBTRCP 1671 bp mRNA linear VRT 17-SEP-1993
 LOCUS African clawed frog beta-TrCP mRNA, complete cds.
 DEFINITION M98268
 ACCESSION M98268.1 GI:295542
 VERSION beta-transducin repeats.
 KEYWORDS Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
 1 (bases 1 to 1671)
 Spevak, W., Keiper, B.D., Stratowa, C. and Castanon, M.J.
 Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein with beta-transducin repeats
 Mol. Cell. Biol. 13 (8), 4953-4966 (1993)
 93330289
 8393141

Original source text: Xenopus laevis (library: S. cerevisiae expression library of X.laevis oocytes) cDNA to mRNA.

Location/Qualifiers
 1..1671
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /cell_type="oocyte"
 /tissue_lib="S. cerevisiae expression library of X.laevis oocytes"
 1..1671
 /gene="beta-TrCp"
 48..1604
 /gene="beta-TrCp"
 /codon_start=1
 /protein_id="AA02810.1"
 /db_xref="GI:295543"
 /translation="MEGFCSLQPTASERDNRDPKILTEKNTLPOTKLANGTSSMIVPQKRLSANVEKEKLVKFKQWSECDQVEVHLIRMKCHYQKHNTYLIKPMRLDFTALPARGLDHIAENILSYLDKSLCSAEVCKEYRVYTSGLMVKLLIERMVRDLSLGRUAEERGGQYLFKNKPPDGKTPPNSFYRALYPKIIDTETIESNRGGRHSQRIHCRSETSGYCYQDDQKIVSGLRDNTIKINDKTLCKRVLMGHTGSVLCLOYDERVITGSSDSYRVMDVNTGMLNLIHCEAVLHLRFNMGMYTCSDRSIAVMDASATDITLRLVLGHRAAVNVDFDKYIVSASGDTIKWNTSCEFRVLNKHGKIACLOYEDRLVSGSDNIRLWDLCEGACIRVLGHEELVRCIRPDNKRIVSGAYDGKKIYMDLVAALDPAPAGTLCRLTIVHSGRVRLQFDEFQIVSSSHDDTILIWFLNDPGLA"

1671
 /gene="beta-TrCp"
 /NT 495 a 341 c 424 g 411 t
 t Scores:
 2.27e-233 Length: 1671
 2597.00 Matches: 488
 92.87% Conservative: 7
 91.56% Mismatches: 8
 85.60% Indels: 30
 5 Gaps: 1

1-168b-2 (1-569) x XELSBTRCP (1-1671)
 18 SerSerGluArgGluAAspCysAsnAsnGlyGluProProArgLysIleIleProGluLys 37
 84 GCCTCAGACAGAGAGACTGTACAGGATGACCGCTAGGAGATTAACCGAGAAA 143
 38 AsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrVal 57
 144 AACCACTTAGACAG----- 158
 58 CysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsn 77
 159 -----ACAAACTTGCCAAAT 173
 78 GlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGlu 97
 174 GGCACCTCCAGCATGATTGTGCCCAACGACGCAAACTGTGAGCAAAATTACGAGAGGAA 233
 98 LysGluLeuCysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheVal 117
 234 AAAGAGCTATCGCTCAAGTATTTTCAGCAGTGGTCCGAGTCCGATCAAGTAGAGTTTGT 293

QY 118 GluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLys 137
 Db 294 GAACACCTGATATCTCGAATGCCCACATATCAGCATGGACATATAACACTTACCTTAAG 353
 QY 138 ProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAla 157
 Db 354 CCAATGTTTAAAGAGGATTTCATTACCGCACTPACCACTCGCGGACTCGATCACATAGCA 413
 QY 158 GluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLys 177
 Db 414 GAAATATATCTTCTATCTGATGCAAGTCAATGTTCTGCAAGAACTGGTATGTAG 473
 QY 178 GluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetVal 197
 Db 474 GAGTGTATCGAGTGACCTCAGATGGAATGCTCTGAAAGAACTCATAGAGCGGATGTC 533
 QY 198 ArgThrAspSerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTrpLeuPhe 217
 Db 534 CGACAGAGATTCTCTTTGGAGAGGACTGGCAGAAAGAGGATGGGGCAATATTGTGTT 593
 QY 218 LysAsnLysProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrPro 237
 Db 594 AAAACAACACTCCAGATGGGAAACGCCACCAAAATTCATTCTACAGACCGCTTTACCCA 653
 QY 238 LysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeu 257
 Db 654 AAAATTATTCAAGACATAGACAAATCGATGCCAACTGGCGTGTGGGAGACACAGCTTA 713
 QY 258 GlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAsp 277
 Db 714 CAAGAATTCACCTGCGGAGTGAAACAAAGCAAGAGGGGTATATCTGTCTCAGTACGATGAT 773
 QY 278 GlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeu 297
 Db 774 CAGAAGATAGTAAGTGGACTCAGATTAACACCATTAAGATCGGATAAGNATCTTTG 833
 QY 298 GluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu 317
 Db 834 GAGTGCAGCGAGTGTGATGGTGCACACTGGGTGAGTCTCTCTGTCGCAATATGATGAG 893
 QY 318 ArgValIleIleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGly 337
 Db 894 AGAGTAATCATTACTGGCTCTTCAGACTCTACCGTCCGGGTGTGGACGTGAACACAGGA 953
 QY 338 GluMetLeuAsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsn 357
 Db 954 GAAATGTTGAACACGCTGATTCACCACTGTGAGGCTGTGCTGCACCTTGAGGTTTAATAT 1013
 QY 358 GlyMetMetValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerPro 377
 Db 1014 GGCATGATGTCACCTGCTCCAAAGATCGTTCCATTGCAAGTTTGGGACATGGCCTCTCG 1073
 QY 378 ThrAspIleThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValValAsp 397
 Db 1074 ACCGATATCACATTACGAGAGATTCTGTAGGCCACCGAGCTGCCGCTAAATGTGGTGAC 1133
 QY 398 PheAspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThr 417
 Db 1134 TTTGATGACAACTATATAGTCTCTGCACTCTGCTGATCGAACAAATAAGGTTTGGACACC 1193
 QY 418 SerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGln 437
 Db 1194 AGTACATGTGAATTTGTGCGGACATTTGAACGGCCCAAGCGTGGTATTGTCATGCTTACAG 1253
 QY 438 TyrArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIle 457
 Db 1254 TATCAGATCGCTTGTGTGAGTGGTCTTCTGTGCAACACATCAGATTGTGGGATATT 1313
 QY 458 GluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArg 477
 Db 1314 GAATCGGTGCATGTTTGGGGTCTTGGAGGACATGAAGAACTGGTCCGGTGCATCCGC 1373
 QY 478 PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeu 497

```
|||||
1374 TTTGATACACAGAGAAATAGTCAGTGGAGCATATACCGGAAAATTAAGTGTGGACCTT 1433
498 ValAlaLeuLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGlu 517
1434 GTCGCTGCTTTGGACCCCGTGGCGCTGCAGGAGACCTGTGTCTCCGGATCTTTGGAG 1493
518 HisSerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHis 537
1494 CATTCGGGCAGAGTTCCTCGGCTGCAGTTTCACGAGTTCACAGATAGTAGAAGCAGCTCTCAC 1553
538 AspAspThrIleLeuIleTyrAspPheLeuAsnAspPro 550
1554 GATGACATATCCCTCATCTGGATTTCTCAACGCCA 1592
(14)
AB033280 2252 bp mRNA linear PRI 14-MAR-2000
ON Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
XN beta-TRCP2 isoform B, complete cds.
AB033280
BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform B.
; Homo sapiens (human)
SM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
XS Koike,J., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M. and
Katoch,M.
Molecular cloning and genomic structure of the betaTRCP2 gene on
chromosome 5q35.1
Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
20160458
D 10694485
DE 2 (bases 1 to 2252)
XS Katoch,M.
Direct Submission
Submitted (05-OCT-1999) Masaru Katoch, National Cancer Center,
Genetics Division; Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
(E-mail:mkatoch@ncc.go.jp, Tel:81-3-3542-2511(ex.4402),
Fax:81-3-3541-2685)
Location/Qualifiers
1.2252
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q35.1"
/tissue_type="fetal lung"
1..2252
/gene="BTRCP2"
157..1746
/locus_tag="BTRCP2"
/codon_start=1
/product="F-box and WD-repeats protein beta-TRCP2 isoform
B"
/protein_id="BAA92330.1"
/db_xref="GI:7209811"
/translation="MEPDSVIEDKTIILMNTSVMEQNEDESPPKNTLWQINSNGTSV
IVSRKRPSEGNVQKEDCIKYFQWSESQDVEFVHLIIRMGCHYQHGHSYLYKPMW
ORDFTALPEQGLDHAENILSYLDARSLCAELVCKEQRVISEGLMKLTERVWR
TDLWKGLSESRGMDVLFKRPPTDGPNSFYSLPKLIQDIETISNWRGCRHNLQ
RIQRENSKGYCLQVDDEKIIISGLRNSIKIWDKTSLECLVLTGHTGSVLQCLQD
BRVITGSDSDTVRVWDNTGEVNTLIHNEAVLHLRFNSGLMTVCSDKRSIAVWDM
ASADITLRLVYVGHRAVNVVDYKIVSASGDTIKVMSSTCEFRVTLNKHGRK
IACQYRDLRVSGSSONTIRLWIEGACILRVLEGEELVRICRPNKRIVSGVDG
KIKYWDLQALDPRAPASTLCRLTLVHSGRFLQDFEQIISSSDDHLLIWDPLN
VPPSAQNETRSPSTYTYISR"
NT 598 a 501 c 558 g 595 t
it Scores:
```

```
Pred. No.: 9,69e-215 Length: 2252
Score: 2400.00 Matches: 451
Percent Similarity: 86.82% Conservative: 43
Best Local Similarity: 79.26% Mismatches: 35
Query Match: 79.10% Indels: 40
DB: Gaps: 4
US-09-601-168B-2 (1-569) x AB033280 (1-2252)
QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 157 ATGGAGCCC---GACTCGGTGATTGAGCACAAGACCATCGAGCTCATGAACACTTCAGTT 213
QY 21 ArgGluAspCysAsnAsnGlyGluProProArgHisIleIleProGluLysAsnSerLeu 40
Db 214 ATGGAAGATCAAAATGAAGATGAGTCCCAAG-----AAAAATACTCTT 258
QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 259 TGGCAG----- 264
QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 265 -----ATAAGTAATGGAACATCA 282
QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db 283 TCTGTGATCGTCTCCAGAAAGAGGCCATCAGAAAGAAACTATCAAAAAGAAAAAGACTTG 342
QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 343 TGTATTAAATATTTTGACCAAGTGGTCTGAATCAGATCAAGTGGAAATTTGTGAACATCTT 402
QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
Db 403 ATTTCCAGAAATGTGTCAATTATTCAGCATGGACATATTAACTCTTACCTCAAGCCCATGTTG 462
QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
Db 463 CAGCGGGACTTTTATACCGCTTTACCGAGCAAGGCTTAGATCATAGCAGAAAAACATT 522
QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
Db 523 CTTTGTACTGGATGCCAGGTCCTCTGTGCGACGAGAGCTGGTATGTAAAGAAATGCCAG 582
QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
Db 583 CGAGTGATCTCAGAAAGAAATGCTTTTGGAAAGAGCTGATTGAACGAATGGTACGCACTGAT 642
QY 201 SerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
Db 643 CCCCTATGGAAGAGACTTTCAGAAAGAGAGGGTGGGATCAGTACCTGTTTAAAAAACAGA 702
QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
Db 703 CCCACAGATGCG-----CCTCCAAATTCATTTTATAGTCAATATATACCAAGATTATC 756
QY 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db 757 CAGGATATAGACACTATAGAATCTTAACCTGGGGGTGTGGACGACACAACTTCAGAGGATT 816
QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
Db 817 CAGTCCCGCTCTGAAATAATAGTAAGTGTCTACTCTTTTACAGTACGATGATGAANAATT 876
QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
Db 877 ATCAGTGGCTACGAGATAATTCATTATTAAGATATGGGATAAACCCAGCCTCGAATGTTG 936
QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 937 AAAGTGTTAAACAGGACACACAGGCTCTGTCTCTCTCTGTCAGTATGATGAGCGTGTCTATT 996
```

321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
:::|||||
997 GTAACCTGGCTCTTACAGATTCACCGGTGAGATGGTGGATGGAACACGGGTGAAGTTCCT 1056
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1057 AACACATTGATCCACCACCAATGAGGCTGTATTGCACATTACGCTTCAGCATGGACTGATG 1116
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1117 GTGACCTGTTCCAGGACCGCTCCATTCCTGTGTGGACATCGCTTCTGGCAGCGACATC 1176
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1177 ACTTTACCGCGTGCTCTGTGTGGCCACGCGGCTGCGCTCAATGTAGACTTTGACGAC 1236
401 LysTyriIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1237 AAGTACATCGTGTCTGCCCTCTGGTGACAGGACCATCAAGTCTGGAGCAGCAGCACCTGT 1296
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1297 GAATTTGTTCTGTACTCTCAATGGGCACCAAGCGGGCAFTGGCTGTCTCCAGTACAGGGAT 1356
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1357 CGCTTGGTGTGTAGTGGATCATCAGATATACCATTAGCTCTGGGATATTGAATGTGGT 1416
461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
1417 GCCTGTTTAAAGAGTCTTAGAGGACATCAAGAAATTTGGTCCGATCATCCGCTTTGATAAC 1476
481 LysArgIleValSerGlyValArgGlyLysIleLysValTrpAspLeuValAlaAla 500
1477 AAGAGATTGTCTAGTGGGCGCTATGATGGGAAATTAAGATTGGGACTTCCAGACTGCT 1536
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1537 CTGACCTCTGAGCGCCAGCAGCAGCATTTGTGTTGGCAGCATTTGGTGGACATCTCTGA 1596
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540
1597 CGTGTGTTTCCGCTCCAGTTTGAATGAGTTTTCAGATCATCAGCAGCTCCCATGTCAGACT 1656
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
1657 ATTTTGAATTTGGATTTCTTAATGTGCTCCAGTCCGACGATGAGACCCGTTCTCCC 1716
561 SerArgThrTyrThrTyrIleSerArg 569
1717 TCCAGAACATACACTTACATCTCTAGA 1743

5

BC026213 4395 bp mRNA linear PRI 08-APR-2002
ON Homo sapiens, F-box and WD-40 domain protein 1B, clone MGC:21122
IMAGE:4419029, mRNA, complete cds.
N BC026213
BC026213.1 GI:20070727
MGC.
Homo sapiens (human)
SM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
E 1 (bases 1 to 4395)
S Strausberg, R.
Direct Submission
L Submitted (02-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk

Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Prepared by: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 27 Row: i Column: 22.

FEATURES

source

Location/Qualifiers

1..4395

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="LocusID:23291"

/db_xref="taxon:9606"

/clone="MGC:21122 IMAGE:4419029"

/issue_type="Lymph, lymphoma"

/clone_lib="NIH_MGC_85"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

372..1961

/codon_start=1

/product="F-box and WD-40 domain protein 1B"

/protein_id="AAH26213.1"

/db_xref="GI:20070728"

/translation="MEPDSVIEDKTIETLMTNTVMEDQNEDESPKKTLLWQISNTSSV
IVSRKREGNYQKEKDLCKYFQWSESDQVEFVHLIRMHYQHGHSINSLKPLM
ORDFTALPEOGLDIAENILSYLDAELSLCAELVCKEQRVISEGMLKKLIERMVR
TDPWKGLSERGMDVLFKNRPTDGPNSFYSLPKIIDIETISNWFGRHNLQ
RIQSESNKGVYCLQYDDEKILISGLRNSIKINDKTSLECLKVLTKHTSGVLCLOVD
ERVITGSSDSTVRVMDVNTGEVLTLLHNEAVLHURFNSGLMVTCSKRSIAWDM
ASATDITLRRVLVGHRAAVNYDFDDKIYVSAGDRTIKVMSTSTCFVRLNGLHKG
IACQLRDLLVVGSSDNTIRLWIECGACLRVLEGHEELVRCIRFONKRVISGAYDG
KIKWDLQALDPRAPASTLCRLTRVHSGRVFLQPDFQIISSHDDTLILWDFLN
VPPSAQNETRSPSRVTYISR"

BASE COUNT 1131 a 928 c 1039 g 1297 t

ORIGIN

Alignment Scores:

Pred. No.:	2,42e-214	Length:	4395
Score:	2400.00	Matches:	451
Percent Similarity:	86.82%	Conservative:	43
Best Local Similarity:	79.28%	Mismatches:	35
Query Match:	79.10%	Indels:	40
DB:	9	Gaps:	4

US-09-601-168B-2 (1-569) x BC026213 (1-4395)

QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 372 ATGGAGCCC---GATTCGGTGTATTGAGGACAGACCATCGAGCTCATGAACACTTCAGTT 428
QY 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
Db 429 ATGGAAGATCAAAATGAAGATGATGATCCCAAG-----AAAAATACTCTT 473
QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 474 TGSCAG-----
QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 480 -----ATAAGTAATGAACATCA 497
QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db 498 TCTGTGATCGTCTCCAGAAAGAGGCCATCAGAGGAAGAACTATCAAAAGAAAAAGACTTG 557

101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 558 TGTATTAAATATTTCGACAGTGGTCTGTAATCAGATCAAGTGGAAATTGTGGACATCTT 617
 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 618 ATTTTCAGCAATGTGTCATTATCAGCATGGACATATTAACTCTTACCTGAAGCCCATGTTG 677
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 678 CAGCGGACITTTATTACCGCTTTTACCAGAGCAAGGCTTAGATCATCATAGCAGAAAAACATT 737
 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr 180
 738 CTTTCGTAACCTGGATGCCAGCTCTCTGTGTGAGCAGAGCTGGTATGTAAAGAAATGGCAG 797
 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 798 CGAGTGATCTCAGAAAGATGCTTTTGGAAAGAGCTGATTGAACGAATGGTACGCACCTGAT 857
 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
 858 CCCCTATGGAAAGACCTTCAGAAAGAAAGAGGCTGGATCATGTAACCTGTTTAAAAACAGA 917
 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 918 CCCACAGATGCG-----CCTCAAAATTCATTTATAGTGTATTATCCCAAGATTATC 971
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 972 CAGATATAGAGACTATAGAAATCTAACTCGCGGTGTGGACAGACAACACTTCACAGAGATT 1031
 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 1032 CAGTGGCGCTCTGAAATATGTAAGGTGTCTACTGTTTACAGTACGATGATGAANAATT 1091
 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 1092 ATCAGTGGCTTACAGATAATTCTATTAAAGATATGGGATAAAACAGCAGCTGGAAATGTTG 1151
 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 1152 AAAGTTTATACAGACACACAGGCTCTGTCTCTGTCTGCAGTATGATGACGGTGTCTTT 1211
 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 1212 GTAACTGGCTCTTCAGATTCTACGGTGGAGATGTGGATGTGAACACGGGTGAAGATCTCT 1271
 341 AsnThrLeuIleHisHisCysGluAlaValIleHisLeuArgPheAsnAsnGlyMetMet 360
 1272 AACACATTGATCCACCAATGAGGCTGATTGCACTTACGGTTCGCAATGTAAGTACTGATG 1331
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1332 GTGACCTGTTCCAAAGGACCGCTCCATTGCTGTGTGGACATCGCTTCTGCGACCGCATC 1391
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 1392 ACTTTACCGCGTGTCTGTTGTGGCCACCGGCGTGGCGTCAATGTASTAGACTTTTGACGAC 1451
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1452 AAGTACATCGTGTCTGCTCTGTGTGACAGGACCCATCAAAAGTCTGGAGCACGCGACCTGT 1511
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 1512 GAAATTTGTTCTACTCTCAATGGGCACAACGGGGGCAATTCCTGTCTCCAGTACAGGGAT 1571
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1572 CCCTCGGTTGTTAGTGGAATCATCAGATAATACCATTAGGCTCTGGGATATTGAATGTGGT 1631
 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480

1632	GCTGTTTAAAGAGTCCTAGAGGACATGAAGAAATTGGTCCGATGATCGCGTTTGATAAC	1691
Db		
Qy		
481	LYsArgILleValSerGLyAlaTVrAspGLyLYsILleLYsValTrpAspLeuValAlaAla	500
Db		
1692	AAAGAGANTTTCAGTGGGGCCCTATGATGGGAAANATAAAGTTTGGGACTTGCAGCTGCT	1751
Qy		
501	LeuAspProArgAlaProAlaGLyThrLeuCYsLeuArgThrLeuValGLuHisSerGLy	520
Db		
1752	CTTGACCTTCGAGCCCCAGCAGCACATTTGTGTTTCCGCACATTTGTTGGAACTTCTGGA	1811
Qy		
521	ArgValPheArgTLeuGLnPhelAspGLuPheGLnILleValSerSerSerHisAspAspThr	540
Db		
1812	CGTGTGTTTCGGGCTCCAGTTTGATGATAGTTTCAGATCATCAGCAGCTCCCATGATGACACT	1871
Qy		
541	ILleLeuILleTrpAspPheLeuAsnAspProAlaAlaGLnAlaGLuProProArgSerPro	560
Db		
1872	ATTTTGNATTTGGGATTTCTTAATATGTGCTTCCAGTCCCGAGAAATGAGACCGGTTCTCCC	1931
Qy		
561	SerArgThrTyrThrTyrIleSerArg	569
1932	TCCAGAACATACACTTACATCTCTAGA	1958
Db		

Search completed: October 22, 2003, 16:26:46
Job time : 5030 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ein - nucleic search, using frame_plus_p2n model

October 22, 2003, 14:49:27 ; Search time 2665 Seconds
(without alignments)
5189.211 Million cell updates/sec

US-09-601-168b-2
score: 3034
e: 1 MPFAEVLQEKALKFMSNSE.....FAAQAEPPSPSRYYTIVISR 569

table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 22781392 seqs, 12152238056 residues
Number of hits satisfying Chosen parameters: 45562784

DB seq length: 0
DB seq length: 2000000000
Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

line parameters:
frame_p2n.model -DRV=xlp
21/USPTO.spool_p/US09601168/runat_22102003_100546_17791/app_query.fasta_1.711
-QFMT=fastap -SUFFIX=st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-N=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-PTO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
309601168 @CGN_1.1 2874 @runat_22102003_100546_17791 -NCPU=6 -ICPU=3
>LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
4EOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
r=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

3: EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssI:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2951	97.3	2897	11	AK032221	Mus muscu
2	2940	96.9	2970	11	AK083257	Mus muscu
3	2686	88.5	4502	11	AK052317	Mus muscu
4	2395	78.9	3824	11	AK041532	Mus muscu
5	1561	51.5	1136	14	CD500760	CD446-H04
6	1407.5	46.4	924	14	CA985560	CA985560 AGENCOURT
7	1390	45.8	895	9	AL898993	AL898993 AL898993
8	1353	44.6	966	14	CA980902	CA980902 AGENCOURT
9	1350	44.5	776	14	CB244526	UI-M-EY0-
10	1348	44.4	810	12	BM944304	UI-M-EH0p
11	1332	43.9	778	13	BQ44188	UI-M-EX0-
12	1323.5	43.6	1002	14	CA971789	AGENCOURT
13	1320.5	43.5	935	14	CA975032	AGENCOURT
14	1310	43.2	850	14	CA327271	UI-M-FY0-
15	1306	43.0	744	14	CB520936	UI-M-GH0-
16	1298.5	42.8	850	13	BQ771102	UI-M-FI0-
17	1298.5	42.8	1201	9	AL556068	AL556068 AL556068
18	1290	42.5	749	14	CB246650	UI-M-FI0-
19	1270	41.9	806	13	BU750387	CH3#032 H
20	1241	40.9	796	14	CA348002	679219 NC
21	1238	40.8	791	14	CA320826	UI-M-FW0-
22	1238	40.8	827	14	CB524867	UI-M-EY0-
23	1227	40.4	804	14	CA790902	AGENCOURT
24	1226.5	40.4	972	13	BQ940482	AGENCOURT
25	1225	40.4	699	14	CB246901	UI-M-FI0-
26	1222.5	40.3	876	13	BU170327	AGENCOURT
27	1222	40.3	732	12	B1088274	602852735
28	1218	40.1	888	12	B1455545	603173974
29	1214	40.0	846	9	AL879775	AL879775 AL879775
30	1184	39.0	1081	12	BM542586	AGENCOURT
31	1184	39.0	1201	9	AL531016	AL531016 AL531016
32	1181	38.9	870	13	BQ463637	603772881
33	1177	38.8	851	13	BQ425358	AGENCOURT
34	1174	38.7	963	13	BX349157	BX349157 BX349157
35	1169	38.5	665	9	AL642623	AL642623 AL642623
36	1169	38.5	758	13	BUG11165	UI-M-FI0-
37	1168	38.5	523	13	BQ913488	AGENCOURT
38	1166	38.4	973	12	BM465280	AGENCOURT
39	1157	38.1	667	9	AL882125	AL882125 AL882125
40	1154	38.0	712	14	CA513313	UI-R-FU0-
41	1149	37.9	985	13	BX419479	BX419479 BX419479
42	1146	37.8	752	9	AU002962	AU002962 AU002962
43	1143	37.7	779	14	CD350839	UI-M-GI0-
44	1143	37.7	699	13	BUG13203	UI-M-EW0-
45	1141	37.6	761	14	CA345696	676407 NC

ALIGNMENTS

RESULT 1
AK032221
LOCUS
DEFINITION
Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430504E23 product:beta-transducin repeat
containing protein, full insert sequence.
ACCESSION
AK032221
VERSION
AK032221.1 GI:26082733
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

CE 1 Carninci, P. and Hayashizaki, Y.
 RS High-efficiency full-length cDNA cloning
 AL Meth. Enzymol. 303, 19-44 (1999)
 NE 99279253
 CE 10349636

CE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 RS Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 AL Normalization and subtraction of cap-trapper-selected cDNAs to
 NE prepare full-length cDNA libraries for rapid discovery of new genes
 CE Genome Res. 10 (10), 1617-1630 (2000)
 AL 20493374
 NE 11042159
 CE 3

CE 4 Kawai, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 RS Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 AL Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 NE Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 CE Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 RS Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 AL Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 NE RIKEN integrated sequence analysis (RISA) system--384-format
 CE sequencing pipeline with 384 multicapillary sequencer
 RS Genome Res. 10 (11), 1757-1771 (2000)
 AL 20530913
 NE 11076861

CE 5 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 RS Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 AL Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 NE Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 CE Kadota, K., Matsuda, H., Ashburner, M., Batalow, S., Casavant, T.,
 RS Fleischmann, W., Gaasterland, T., Gassi, C., King, B., Kochiwa, H.,
 AL Kuehl, P., Lewis, S., Matsumoto, Y., Nikaide, I., Pesole, G.,
 NE Quackenbush, J., Schriml, L. M., Stauff, F., Suzuki, R., Tomita, M.,
 CE Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 RS Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 AL Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C.,
 NE Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 CE Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 RS Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 AL Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 NE Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 CE Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
 RS Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
 AL and Hayashizaki, Y.
 NE Functional annotation of a full-length mouse cDNA collection
 CE Nature 409 (6821), 685-690 (2001)
 RS 21085660
 AL 11217851

CE 6 The FANTOM Consortium and the RIKEN Genome Exploration Research
 RS Group Phase I & II Team.
 AL Analysis of the mouse transcriptome based on functional annotation
 NE of 60,770 full-length cDNAs
 CE Nature 420, 563-573 (2002)
 RS 5 (bases 1 to 2897)

CE 7 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 RS Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 AL Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 NE Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 CE Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 RS Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 AL Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 NE Okazaki, Y., Saito, H., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 CE Sano, H., Sasaki, D., Shibata, K., Shingagawa, A., Shiraki, T.,
 RS Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 AL Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 NE Muramatsu, M. and Hayashizaki, Y.
 CE Direct Submission

AL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 RS Physical and Chemical Research (RIKEN), Laboratory for Genome

COMMENT
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

FEATURES
 source
 1..2897
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:6430504E23"
 /db_xref="taxon:10090"
 /clone="6430504E23"
 /sex="male"
 /tissue_type="olfactory brain"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 misc_feature
 1..2897
 /note="beta-transducin repeat containing protein
 (MGD|MG1:1338871, GB|NM_009771, evidence: BLASTN, 100%,
 matches=1502)"

BASE COUNT 720 a 736 c 771 g 670 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1..376-314 Length: 2897
 Score: 2951.00 Matches: 561
 Percent Similarity: 93.07% Conservative: 3
 Best Local Similarity: 92.57% Mismatches: 57
 Query Match: 97.26% Indels: 37
 DB: 11 Gaps: 1

US-09-601-168B-2 (1-569) x AK032221 (1-2897)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
 Db 68 ATGACCCCGCAGAGCGGTGCTGCAGAGAAAGCGCTTAAGTTATGTGCTCTATGCC 127
 Qy 16 ----- 16
 Db 128 AGGTCTCTGTGGTGGGCTGCTCCAGCCTGGCGAGCAGCATGCCCTTCGCTGCATGCC 187
 Qy 17 -----AsnSerSerGluArgGluAspCy 24
 Db 188 TATAACCCAGGCACTGGCGCACTCAGAGCTTTCAGAGAAATTCCTCAGAGAGAGAGACTG 247
 Qy 24 sAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeuArgGlnThrTy 44
 Db 248 TATAATGGCGAACCCCTTAGAGAGTAATACAGAGAAAGATTCACCTTAGACAGACTTA 307
 Qy 44 rAsnSerCysAlaAArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMe 64
 Db 308 CAACAGCTGGCCAGGCTTTGCATAAACCAAGAGCAGTATGCTTAACAAGCAGCTGCTAT 367
 Qy 64 tlyThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVa 84
 Db 368 GAAGACTGAAATTTGTGTGGCAAAAGCCAAATTCGCAATGGCCTTCCAGCATGATTGT 427
 Qy 84 lProLysGlnArqLysLeuSerAlaSerTyrcGluLysGluLysGluLeuLysVallysty 104
 Db 428 GCCCAAGCAGCGGAAACTCTCAGCAAGCTATGAGAGAGAAAGGAGAGCTGTGTCAAGTA 487
 Qy 104 rPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSerGlnMe 124
 Db 488 TTTTGGCAGTGGTCAGACTCTGATCAAGTGGAAATTTGTAGAACACCTTATATCCAAAT 547

```

124 tCysHisTyrGlnHisGlyHisIleAsnSerTyrIleuLysProMetLeuGlnArgAspPh 144
548 GTGTCACTACCAAGCATGGGCACATCAACTCTCTACCTAAACCTATGCTGCAGAGGGATT 607
144 eileThrAlaLeuProAlaAArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLe 164
608 CATAACTGCACCTGCCAGACAGGGGCTCTGGACCAATCGCTGAGAACATTTCTGTCAATCT 667
164 uAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSe 184
668 GGACGCCAAGTCACTGTGTGCTGTGCTGAGCTGTGCAAGGAATGTTACCGCTGACGTC 727
184 rAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpAr 204
728 GGACGGCATGCTGTGGAANAAGCTCATCGAGGAGTGTGAGGACGACTCTGTGTGGCG 787
204 gGlyLeuAlaGluAArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspG1 224
788 AGGCTTGGCAGAGCGCAGAGGCTGGGACAGTACTTATCAAAAACAAACCTCTCATGA 847
224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG1 244
848 GAACGCTCTCCCACTCTCTTTATAGAGCGCTTTATCTCTAAATCATACAAGACATTGA 907
244 uThrIleGluSerAsnTrpAtGysGlyAArgHisSerLeuGlnArgIleHisCysArgSe 264
908 GACAAATAGAGTTCNAATTGGAGATGTGGGCGACATAGTTTACAGAGATCCACTGCCGGAG 967
264 rGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLe 284
968 TGAACAACAGTAAGGGGTTTACTGTGTTACAGTACGACACAGACAGATAGTCAGCGGCT 1027
284 uArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuTh 304
1028 TGAGACAACACCATCAAGATCTGGATTAAGACACACTGGATGCAAGCGGATCTCAC 1087
304 rGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySe 324
1088 GGGCCACACGGGCTCGTCTCTGCTGCTGAGTACGATGAGAGGAGTATCATCAGAGGCTC 1147
324 rSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuI1 344
1148 CTCAGACTCCACCGTCAGAGTGGGATGTAAATCCAGTGGAGATGCTTAACACATTTGAT 1207
344 ehHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSe 364
1208 TCACCACTGTGAAGCGTTCTGCACCTGGCTTCAATAATGGCATGATGGTGACCTGTC 1267
364 rLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgAr 384
1268 CAAGACCGCTTCATCGCTGTGTGGGATATGGCTTCCGCCAATGACATCACCCCTCAGGAG 1327
384 qValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleVa 404
1328 GTGTGCTGTGGGACACCGAGCTGGCTCAATGTTGATGCTTGTGATGCAATGATCATGCT 1387
404 lSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValAr 424
1388 TTCTGCTCTGGAGATAGAACCAATAAAGGTGTGGAACACACAGTACCTGTGAATTCGTAAG 1447
424 gThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVa 444
1448 GACCCCTAAATGGGCACAAGCGTGGCATCGCTGTTGCGATGACAGACAGCGCTGGTGGT 1507
444 lSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuAr 464
1508 GACGGCTCTCTGCACAAACCAATCAGGCTGTGGGCATAGAGTGTGGAGCATGCTCGC 1567
464 gValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleVa 484
1568 AGTGTGGAGGGCCATGAGGAGTTGTCACGCTGCATTCGATTTGATTAACAAAGATAGT 1627

```

484 lSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProAr 504
1628 GAGCGAGCCCTATGATGGGAAATTAAGTGTGGATCTTATGGCTGCTTTGGACCGCG 1687
504 gAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheAr 524
1688 TGCTCCACAGGGACTCTCTGTCTGCGACACTTGTGGAGCATCTCGAAGAGATTTCCG 1747
524 gLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThrIleLeuIleTr 544
1748 CTCCTCAGTTTATGAAATTCAGATTGTCTAGTAGTTCATGATGACACAAATTCATCTG 1807
544 pAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerProSerArgThrTy 564
1808 GGACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867
564 rThrTyrIleSerArg 569
1868 CACCTACATCTCCAGA 1883

RESULT 2
AK083257
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

AK083257 2970 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:C630031K01 product:beta-transducin repeat
containing protein, full insert sequence.
AK083257
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, P., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzei, J., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2970)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

Location/Qualifiers

1..2970

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:C630031K01"

/db_xref="taxon:10090"

/clone="C630031K01"

/sex="male"

/tissue_type="hippocampus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

81..1900

sc_feature

/note="beta-transducin repeat containing protein (MGDI|MG1:1338871, GB|NM_009771, evidence: BLASTN, 100%, match=1502) putative"

UNT 739 a 758 C 788 g 685 t

nt Scores:

O.: 2.35e-313 Length: 2970
2940.00 Matches: 560
Similarity: 92.92% Conservative: 4
cal Similarity: 92.26% Mismatches: 5
match: 96.90% Indels: 38
11 Gaps: 1

US-09-601-168B-2 (1-569) x AK083257 (1-2970)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
Db 81 ATGACCCGGCAGAGCGGGTGTGTCAGAGAGAAAGCGCTTAAGTTTATGTGCTTATGCC 140
Qy 16 ----- 16
Db 141 AGGTCTCTGTGGCTGGGTCTCTCAGCCTGGCGGACAGCATGCTTCTGGTGGATGCCTG 200
Qy 17 -----AsnSerSerGluArgGluAspCys 24
Db 201 TATAACCCAGGAGACTGGGCACCTCAAGCTTCCAGAAATTCCTCAGAGAGAGAACTGT 260
Qy 25 AsnAsnGlyGluProArgLysIlelleProGluLysAsnSerLeuArgGlnThrTyr 44
Db 261 AATAATGGCGAACCCTTAGGAAGATAAATACCAGAGAGAAATTCATTAGACAGACTTAC 320
Qy 45 AsnSerCysAla-ArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMe 64
Db 321 AACAGCTGTGCCAGGCTTTGCATAAACCAAGAGACAGATATGTCTAACAGCACTGCTAT 380
Qy 64 tLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVa 84
Db 381 GAAGACTGAAATTTGTGGCCAAAGCCAAATTCGCAATGGCACTTCCAGCAATGATTGT 440
Qy 84 lProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLysValLysTy 104
Db 441 GCCCAAGCAGCGGAAACTCTCAGCAAGCTATGAGAGAGAAAGAGACCTGTGTGTCAGTA 500
Qy 104 rPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisIleSerGlnMe 124
Db 501 TTTTGAGCAGTGTGTGAGAGTCTCAAGTGGAAATTTAGAACACCTTATATCCCAAT 560
Qy 124 tCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPh 144
Db 561 GTCTCACTACCAGCATGGGCACATCACTCTACCTAAACCTATGCTGCAGAGGGATT 620
Qy 144 elleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrIe 164
Db 621 CATAACTGCACCTCCAGCAGCGGGTGTGGACACATCGCTGAGAACATTCGTCTCATAC 680
Qy 164 uAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSe 194
Db 681 GGACGCCAAGTCACTGTGTCTGTCTGAGTGTGTCGAAAGAAATGTACCGGTGAGTC 740
Qy 184 rAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpAr 204
Db 741 GGACGCCATGCTGTGGAAAAAGCTCATCGAGAGAGTGGTCAGGACGGACTCTCTGTGG 800
Qy 204 gGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProAspGly 224
Db 801 AGGCCTGGCAGAGCGGAGGCTGGGACAGTACTTATTCAAAAACAAACCTCTCTGATGA 860
Qy 224 YAsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG 244
Db 861 GAACGCTCTCTCCCACTCTTTTATAGAGCGCTTTATCTTAAATCATACAGACATTGA 920
Qy 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
Db 921 GACAAATAGATGCCAATTGGAGATGTGGCGACATAGTTTACAGAGAAATCCACTGCCGG 980
Qy 264 rGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLe 284
Db 981 TGAACAAGATAAAGGGGTTTACTGTGTTTACAGTACGACGACCCAGAGATAGTCAGCG 1040
Qy 284 uArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuTh 304
Db 1041 TCGAGACAAACCATCAAGATCTGGGATAAAGCACACCTGGAATGCAAGCGGATTCTCAC 1100
Qy 304 rGlyHisThrGlySerVal-LeuCysLeuGlnTyrAspGluArgValIleIleThrGlyS 324

1101 GGGCCACACGGGCTCCGTCCTGTCCTGTCAGTACGATGAGAGGGTGATCATCAGGCT 1160
324 erSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuI 344
1161 CCTCAGACTCCACGTCAGAGTGGGATGTAAATGACAGGTGAGATGCTAAACACATGA 1220
344 leHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCys 364
1221 TTCACCACTGTGAAGCCGCTTCGACCTGCGCTTCAATATGGCATGATGGTGACCTGTT 1280
364 erLysAspArgSerLeuAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgA 384
1281 CCAAGACCGTTTCCATCCTGCTGGGATTTGGCTTCCCACTGACATCACCTCAGGA 1340
384 rgValLeuValGlyHisArgAlaAlaValAsnValAlaPheAspAspLysTyrlleV 404
1341 GGGTGTGGTGGACACCGAGCTCGGTCATGTTGTAGATTTTGATGACAAAGTACATCG 1400
404 alSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValA 424
1401 TTTCTGCTCTGGAGATAGAACCATAAAGGTGTGGAACACAAAGTACCTGTGAATTCGTAA 1460
424 rgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyArgAspArgLeuValV 444
1461 GGACCCCTAAATGGGCACCAAGCGTGGCATCGCTGTTTGCAGTACAGAGACAGGCTGGT 1520
444 alSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuA 464
1521 TGAGCGGCTCTCTGACAAACCATCAGGCTGTGGACATAGATGTGGAGCATGCTCTGC 1580
464 rgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsnLysArgIleV 484
1581 GAGTGTGGAGGGCATGAGGAGTGTGACGCTGATTCGATTCGATTTGATAACAAAGGATAG 1640
484 alSerGlyAlaTyAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProA 504
1641 TGACGGAGGCTATGATGGGAAATTAAGTGGGATCTTATGGCTTTTGACCCGC 1700
504 rgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheA 524
1701 GTGCTCCAGCAGGACTCTCTGCTCGGACACTTGTGGAGCATCTTGGAGAGTTCCT 1760
524 rgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleT 544
1761 GCCTCCAGCTTTCATGAATTCAGATTTGCAGTAGTTCACATGATGACACAAATTCATCT 1820
544 rpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerProSerArgThrT 564
1821 GGGACTTCTGAATGATCCAGCTGCTCAGCTGAACCCGCCCTCTCCCTTCCTCGGACAT 1880
564 yrThrTyrlleSerArg 569
1881 ACACCTACATCTCCAGA 1897
AK052317 4502 bp mRNA linear HTC 05-DEC-2002
Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330027D11 product:beta-transducin repeat containing protein, full insert sequence.
AK052317
AK052317.1 GI:26095121
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gaibolli, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, F., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
21085660
11217851
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
12466851
6 (bases 1 to 4502)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Scgabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

Location/Qualifiers
1. 4502
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D330027D11"
/db_xref="taxon:10090"
/clone="D330027D11"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="13 days embryo"
1. 4502
/note="beta-transducin repeat containing protein
(MGD|GI:1338871, GB|NM_009771, evidence: BLASTN, 100%,
match=1502)"

JNT 1284 a 998 c 1126 g 1094 t

nt Scores:

5.72e-285 Length: 4502
2686.00 Matches: 509
90.46% Conservative: 3
89.93% Mismatches: 4
88.53% Indels: 50
11 Gaps: 1

01-168B-2 (1-569) x AK052317 (1-4502)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
88 ATGGACCCGGCAGAGCGGTCTGCGAGAGAGAGCGCTTAAGTTTATGGCTGCAGTCCG 147
16 ----- 16
148 GCTGGTGATATTGGAGGAATAATAAAGTGCTCTATGCCAGGTCTCTGGCTGGC 207
16 ----- 16
208 TGCTCCAGCTGGCGGACAGCATGCTTCGCTGGCATGCCCTGTATAACCCAGGAGCTGC 267
17 -----AsnSerSerGluArgGluAspCysAsnAsnGlyGluProPro 30
268 GCACTCACAGCTTCCAGAAATTCCTCAGAGAGAGAACTGTATAATATGGCGAACCCCT 327
31 ArgLysIleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeu 50
328 AGGAAGATAAATACCAAGAGAAAGAGCTGTGTGCAAGTATTATGACAGCTGTGCCAGCT 387
51 CysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysVal 70
388 TGCATAAACCAAGAGACAGTATGCTTAAACAGCACTGTCTATCAAGACTGAAATTTGTGT 447
71 AlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeu 90
448 GCCAAGCCAAACTTGGCAATGGCACTTCCAGCATGATTGTGCCAAGCAGCGGAATC 507
91 SerAlaSerTyrGluLysGluLysGluLeuCysValLysTyrPheGluGlnTrpSerGlu 110
508 TCAGCAAGCTATGAGAAGAAAAGAGCTGTGTGCAAGTATTATGACAGCTGTGTCAGAG 567
111 SerAspGlnValGluPheValGluHisLeulleSerGlnMetCysHisTyrGlnHisGly 130
568 TCTGATCAAGTGGAAATTTGTAGAACACCTTATATCCAAATGTGTCACTACAGCATGG 627
131 HisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAla 150

628 CACATCAACTCTACCTAAAACTATGCTGCAGAGGGATTTTCATAACTGCACCTGCAGCA 687
151 ArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCys 170
688 CGGGGTCTGGACCACATCGCTGAGAACAATTCCTGATACCTTGTGGACGCAAGTCACTGTG 747
171 AlaAlaGluLeuValCysLysGluTrpTyrArgValThrSerAspGlyMetLeuTrpLys 190
748 GCTGCTGAGCTCGTGTGCAAGGAATGGTACCGCGTGACGTGGACGGCATGCTGTGGA 807
191 LysLeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGluArg 210
808 AAGCTCATCGAGAGGATGCTCAGGACGGACTCTCTGTGGCAGGCTCTGCAGAGCGCAGA 867
211 GlyTyrGlyGlnTyrLeuPheLysAsnLysProAspGlyAsnAlaProProAsnSer 230
868 GGCTGGGACAGTACTTATTCAAACCAAACTCTCTGATGAGAACGCTCTCCCACTCC 927
231 PheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGluSerAsnTrp 250
928 TTTTATAGAGCGCTTATCCTAAATCATACAAGACATTGAGACAATAGAGTCCAATTGG 987
251 ArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyVal 270
988 AGATGGGCGACATAGTTTACAGAGAATCCACTCCGGAGTGAACAAAGTAAAGGGGTT 1047
271 TyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLys 290
1048 TACTGTTTACAGTACGACGACCCAGAGATAGTCACGGGCTTCGAGACAACACCATCAAG 1107
291 IleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerVal 310
1108 ATCTGGGATAAAAGCACACTGGAATGCAAGCGGATCTCACGGGCCACACGGGCTCCGTC 1167
311 LeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThrValArg 330
1168 CTGTGCTCCAGTACGATGAGAGGGTGATCATCAGAGCTCTCAGACTCCACCGTCAGA 1227
331 ValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaVal 350
1228 GTGTGGGATGTAATGTCAGGTGAGATGCTTAAACACATTGATTCACCACTGTGAAGCCGTT 1287
351 LeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAla 370
1288 CTGCACCTCGCTTCAATAATGGCATGATGGTGGACCTGTCCAAAGACCGTTCCATCGCT 1347
371 ValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGlyHisArg 390
1348 GTGTGGGATATGGCTTCCCAACTGACATCACCTCAGAGGGGTGCTGGTGGGACACCGA 1407
391 AlaAlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSerGlyAspArg 410
1408 GCTGCGGTCAATGTTGTAGACTTTGATGACAAGTACATCGTTCTCGCTCTGAGATAGA 1467
411 ThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLys 430
1468 ACCATAAAGGTGGGAACAAGTACCTGTGAATTCGTAAAGACCTTAATGGGCAAG 1527
431 ArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGlySerSerAspAsn 450
1528 CGTGGCATCGCTGTTTGCAGTACAGACAGCGCTGGTGGTGGCGGCTCCCTCTGACAAC 1587
451 ThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGlu 470
1588 ACCATCAGCTGTGGACATAGAGTGTGGAGCATGCTCGAGTGTGTGAGGGGCCATGAG 1647
471 GluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyValaTyrAspGly 490
1648 GAGTTGGTACGCTGCATTTCGATTTGATACAAAGAGTAGTGAGCGGAGCCCTATGATGG 1707
491 LysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeu 510

1708 AAAATTAAGTGTGGGATCTTATGCTGCTTTGGACCGCGTGCTCCAGCAGGACTCTC 1767

511 CysLeuArgThrLeuVal 516

||||| ||||| ||||| ||||| |||||

1768 TGTCGCGACACTTGTG 1785

AK041532 3924 bp mRNA linear HTC 05-DEC-2002
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
library, clone:A630019L11 product:F-BOX/WD-REPEAT PROTEIN 1B (F-BOX
AND WD-REPEATS PROTEIN BETA-TRCP2) homolog [Homo sapiens], full
insert sequence.

ON AK041532

AK041532.1 GI:26334548

HTC, CAP trapper.

S Mus musculus (house mouse)

ISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

NE 99279253

ED 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

AL 20499374

ED 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

NE 20530913

ED 11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,

Quackenbush, J., Schriml, L. M., Staab, F., Suzuki, R., Tomita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D. A., Kamly, M., Lee, N. H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodrigues, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,

Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,

Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.

and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

AL NE 21085660

ED 11217851

TITLE

JOURNAL

REFERENCE

AUTHORS

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3924)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, N.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,

Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,

URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,

Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics

Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome

Trust/MRC building Addenbrookes Hospital Cambridge) whose

assistance we gratefully acknowledge

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1. 3824

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:A630019L11"

/db_xref="taxon:10090"

/clone="A630019L11"

/tissue type="thymus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="3 days neonate"

169. 1758

/notes="unnamed protein product; F-BOX/WD-REPEAT PROTEIN 1B

(F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2) homolog [Homo

sapiens] (SWISSPROT|O9UKB1, evidence: FASTY, 99.2%ID,

100%length, match=1587)

putative"

/codon_start=1

/protein_id="BAC30975.1"

/db_xref="GI:26334549"

/translation="MEPDSVIEDKTIELMNTSMVEDQNEDESPKKSALMQISNGTSSV

IYSRKEPSGNYQKRCGLGICVFDQMSDQVEFVEHLISRMCHYOHGHINSYKPLML

QRDTFALPEQGLDHTAENILSYLDARSALCAELCKEQRVISEGMLWKLEIYR

TDPLMKGLSERRGWDQYLFPNRTDGPNSFYRSYFKPKIIDITIESNWRCHNLQ

RIQVSGSSDYCYQDDDKIISGLRDSIKWDKSSLECLKVLGTGSGVLCLOYD

ERVITSGSDSTVRVNDVMTGVLANTLIHNEAVLHFLRFSNGLMTVCSKDRSLAVWDM

AGATDTTLRLVLGHRAAVNVVDFDKYIVSAGDRTIKVSTSTCEPVTNLGHKRG

KIACLOVDRLLVSGSDNTIRLWDLCEGACLEVLGHEELVRCIFRDKRIYSGVDG

KIKVLDQLAALDPAPASTLCRLTVEHSGRVRLQDFDFQIISSHDDTLIWDPLFN

VPPSAQNTRSPSRITYISR"

/note="putative"

3805. 3810

/note="putative"

3824

/note="putative"

885 a 915 c 982 g 1042 t

polyA_signal

polyA_site

BASE COUNT

ORIGIN

The FANTOM Consortium and the RIKEN Genome Exploration Research

at Scores:

6.77e-253 Length: 3824
 2395.00 Matches: 449
 Similarity: 86.47% Conservative: 43
 Mismatches: 37
 Indels: 40
 Gaps: 3

01-168B-2 (1-569) x AK041532 (1-3824)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 169 ATGGAGCC---GACTCGGTGATTGAAGACCAAGACCATCGAGCTCATGAACACTTCAGTT 225
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysLlelleProGluLysAsnSerLeu 40
 226 ATGGAAGATCAAAATGAAGATGAGTCCCAAGAAA----- 261
 41 ArgGlnThrThrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 261 ----- 261
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 262 -----AGTGTCTTTGGCAGATCAGTAATGGAACGTCA 294
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrgluLysGluLysGluLeu 100
 295 TCTGTGATGTCTCAGAAAGAGCGCTCAGAGGGAACTACCAGAAAGAAAGACCTTG 354
 101 CysValLysTyrgluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
 355 TGCATTAAAGTACTTTGACCAGTGTCTGAATCAGATCAGGTGGAAATTTGTGGACATCTT 414
 121 IleSerGlnMetCysHisTyrglnHisGlyHisLleAsnSerTyrgluLysProMetLeu 140
 415 ATCTCAGCGATGTGCATTATCAGATGAGCATATTAATCTTACCTGAAGCCCATGTTG 474
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 475 CAGCGGGACTTATACCTGCTTTACAGAGCAGAGCTTAGATCAGATGAGAAACATT 534
 161 LeuSerTyrgluLeuAspAlaSerLeuCysAlaAlaGluLeuValCysLysGluThrPyr 180
 535 CTCCTCCTAGTGCAGAGTCTCTGTGTGTCAGCAGAGCTGGTGTGTAAAGAAATGGCAG 594
 181 ArgValThrSerAspGlyMetLeuThrLysLysLeuIleGluArgMetValArgThrAsp 200
 595 CGAGTGATCTCAGAGGGATGCTTTGGAGAAAGCTGATTGAGAGGATGGTGGCAGCCGAC 654
 201 SerLeuThrArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrgluLysAsnLys 220
 655 CCTCTCTGGAAGGACTCTCAGAAAGAGAGCTGGGATCAGTACCTGTTTAAAGACAGA 714
 221 ProProAspGlyAsnAlaProProAsnSerPheTyrgluAlaLeuTyrgluLysIlelle 240
 715 CCTACAGATGGC-----CCTCCCAACTCATTATATAGATCATTATACCCAAAGATTATC 768
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGluArgIle 260
 769 CAGGACATAGACCATAGATCAACTGGCGGTGGAGCAGACACAACTTCAGAGGATC 828
 261 HisCysArgSerGluThrSerLysGlyValTyrgluCysLeuGlnTyrgluAspGlnLysIle 280
 829 CAGTGGCGCTCTGAAATAGTAAGGTGTCTACTCTTTGCAATATGATGATGACAAAATT 888
 281 ValSerGlyLeuArgAspAsnThrIleLysLleThrAspLysAsnThrLeuGluCysLys 300
 889 ATCAGTGGCGCTCCGGGACAACTCTATCAAGATCTGGGATAAAGCAGCTTGAATGTTG 948
 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrgluArgValIle 320

949 AAAGTGCTAAACGGGCCACACAGGCTCTGTCTCTCCAGTATGATCAGCGAGTCATT 1008
 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyCysLeuMetLeu 340
 1009 GTAACCTGGTCTTTCAGACTCCACGGTGTAGAGTCTGGGATGTGAACACTGGTGAGTCTC 1068
 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 1069 AACACACTCATCCACCAATGAAGCCGTACTGCACTTACGCTTCAGCAATGGACTGATG 1128
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1129 GTGACTTGTTCACAGGACCGTTCCATTCGCGTGTGGGACATGGCTTCTGCCACCGATATC 1188
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 1189 ACTTTACGCGGTGTCTGTGTGGCCACCGTGTCTGTCATGTAGTAGACTTTTGATGAT 1248
 401 LysTyrlleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1249 AATACATCGTGTCTCTTCAGGACAGGACCAATTAAGTGTGGAGCAGCAGACATGT 1308
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrgAsp 440
 1309 GAGTTGTGTCGCACTCTGAATGGGCACCAAGCGAGGATCGGCTGTCTGCAGTACCGGAC 1368
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1369 CGGCTGTGTGTAGTGATCATCATTAATACCATCCGTTATGGATATTGAATGTGT 1428
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1429 GCCTGTTTAAAGAGTCTCAGAGGGCAGCAAGAAATTTGGTCCGTGCATCCGTTTGAATAAC 1488
 481 LysArgIleValSerGlyValArgGlyLysLysValTrpAspLeuValAlaAla 500
 1489 AAGAGATTGTGATGGCCCTTATGATGGGAAGATTAAAGTCTGGACTTGCAGCTGCT 1548
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1549 CTTGACCTCGGGCCCCCAGCAAGACATTGTCTGTGCACCTTGTGGACACTCTGGA 1608
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540
 1609 CBTGTCTTTCCGCTGCGATTTGATGAGTTTCAGATCATCAGCAGCTCCCATGATGACACT 1668
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1669 ATTTGATTTGGATTCTTAATGTGCTCCAGTGGCCCAATGACACCGCTCTCC 1728
 561 SerArgThrTyrgluSerArg 569
 1729 TCCAGAACCTACACCTATATCTCCAGA 1755

RESULT 5

CD500760 1136 bp mRNA linear EST 12-JUN-2003
 CDA46-H04.x1d-t SHGC-CDA Gasterosteus aculeatus cDNA clone
 LOCUS CDA46-H04 5', mRNA sequence.

DEFINITION

CD500760

ACCESSION

CD500760.1 GI:31427791

VERSION

EST.

KEYWORDS

Gasterosteus aculeatus (three spined stickleback)

SOURCE

Gasterosteus aculeatus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;

Gasterosteidae; Gasterosteus.

1 (bases 1 to 1136)

Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,

Schmutz,J. and Myers,R.M.

Expressed sequence tags from Gasterosteus aculeatus

Unpublished

Contact: Kingsley, DM
 HHMI and Department of Developmental Biology
 Stanford University School of Medicine
 Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
 Tel: 650 725 5954
 Fax: 650 725 7739
 Email: kingsley@cmmg.stanford.edu
 Plate: 46
 High quality sequence stop: 862.

Location/Qualifiers

1. .1136
 /organism="Gasterosteus aculeatus"
 /mol_type="mRNA"
 /strain="Salinas river, CA"
 /db_xref="taxon:69293"
 /clone="CDA46-H04"
 /sex="mixed male and female"
 /tissue_type="heads and internal organs combined"
 /dev_stage="adult"
 /clone_lib="SHGC-CDA"
 /note="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI (5' adaptor); Site 2: XhoI (3' linker primer). The mixed organ cDNA library was generated using the ZAP-cDNA method by Stratagene. First strand cDNA synthesis was primed with a 50 bp linker primer containing an oligo dT sequence preceded by a synthetic XhoI site. 5 prime adaptors were used containing an EcoRI cohesive end. The finished cDNAs were inserted in to the ZAP express vector unidirectionally in the sense orientation with respect to the lacZ promoter of pBK-CMV. An amplified library was prepared from approximately 3 million primary clones in the lambda ZAP Express vector. In vivo excision was then used to generate individual pBK-CMV phagemid clones for EST sequencing."

UNT 262 a 304 c 302 g 266 t 2 others

nt Scores:

1.69e-161 Length: 1136
 1561.00 Matches: 301
 Similarity: 92.98% Conservative: 17
 cal Similarity: 88.01% Mismatches: 23
 Indels: 51.45%
 Gaps: 14

01-168B-2 (1-569) x CD500760 (1-1136)

229 AsnSerPheTyArgAlaLeuTyrProIysIleIleGlnAspIleGluThrIleGluSer 248
 3 AACTCTACTACCACTCCCTTTATCCCAAGATCATCCCAAGACATAGAGACTATTGAGGCA 62
 249 AsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLys 268
 63 AATTGGCGGTGTGGCAGACACAACTTTCAGCGGATTCAGTGTCTCAGAAACAGTAA 122
 269 GlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThr 288
 123 GGGGTTTACTGTCTCCAGTACGATGACGACAAAGATCATCAGTGGGCTTAGGACAAATCC 182
 289 IleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGly 308
 183 ATCAAGATCTGGGATTAAGCAGTCTCTGGAGTGTCTGAAGATATCTACCGGTACACGGGA 242
 309 SerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThr 328
 243 TCAGTGTGTGTCTCAGATGACGAGAGGGTTCATCGTCACCGGTCTCTGACTCGACT 302
 329 ValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGlu 348
 303 GTGAGGGTGTGGAGGTCTACTACGGGTGAGGTACTGAACACGCTGATCCACCAACAGAG 362
 349 AlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSer 368

Db 363 GCGGTTCTTCACTACGCTTGCCTCAACGGGATTGATGGTCACCTGTCTCAAGGACCGCTCG 422
 Qy 369 IleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGly 388
 Db 423 ATCGCGTCTGGGACATGGCTCCCGACTGATCATCGCTACCGCGCTCTCTCGGGGA 482
 Qy 389 HisArgAlaAlaValAsnValValAspPheAspLysIleValSerAlaSerGly 408
 Db 483 CACCGGGTCTGTCAATGTGTGTCGACTTTGATGACAAATACATTTGTCTCGGCTCAGGG 542
 Qy 409 AspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsnGly 428
 Db 543 GACCGTACCAATAAGGTATGGAGCAACAGTACCTGTGAGTTGTGGCCACACTAAATGGT 602
 Qy 429 HisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGlySerSer 448
 Db 603 CATAAGAGAGGATCGCTGTCTACAGTACAGAGATCGCTGGTGTGAGGCTCATCG 662
 Qy 449 AspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGly 468
 Db 663 GACACACAAATCCGTTATGGGATATAGAGTGTGGGGCGTGTTCGGAGTGTCTGGAAGGT 722
 Qy 469 HisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyAlaTyr 488
 Db 723 CACGAGAGTGTGTGCTGCTGCTTTCATTCGCTTTCACATTAAGAGGATTCAGCGGTGCTAC 782
 Qy 489 AspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGly 508
 Db 783 GAGGCGAGATCAGGTGTGGGACCTTCGANGAGCCCTCGA-CCAGAGCCCCCAGCCAGC 841
 Qy 509 ThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArgLeuGlnPheAsp 528
 Db 842 ACATTATGTCTGGCCTCTTAGTGGAGCACTCTGGGGCGGTGTTCGCTCTCCAGTTTGTAT 901
 Qy 529 GluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrpAspPheLeuAsn 548
 Db 902 GAGTTTCAGATCATCAGCAGTCTTCAGGAGTACCATCTCTGAICTGGGACTTTCCTGAC 961
 Qy 549 ---AspProAlaAlaGlnAlaGluProProArgSerProSerArgThrTyrThrIle 567
 Db 962 GTCTCGACTAATGGCAGTGGG-AGGAACGGTCTCCCTCCCGCACTACACTTACATT 1020
 Qy 568 SerArg 569
 Db 1021 TTTAGA 1026

RESULT 6
 CA985560 924 bp mRNA linear EST 06-JAN-2003
 LOCUS AGENCOURT 11076787 NICHG_XGC_001 Xenopus laevis cDNA clone
 DEFINITION IMAGE:6860098 5', mRNA sequence.
 CA985560
 VERSION CA985560.1 GI:27518227
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 924)
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 TITLE Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-f@mail.nih.gov
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM14473 row: 0 column: 09

High quality sequence start: 16

High quality sequence stop: 717.

Location/Qualifiers

1. 924

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clones="IMAGE:6860098"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NICHD XGC OOI"

/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;

Cloned unidirectionally. Primer: Oligo dT. Average insert

size 2.2 Kb. Constructed by Life Technologies."

236 a 188 c 264 g 236 t

UNT

nt Scores:

10.: 1.16e-144 Length: 924

1407.50 Matches: 278

Similarity: 96.22% Conservative: 2

cal Similarity: 95.53% Mismatches: 5

atc: 46.39% Indels: 6

14 Gaps: 1

01-168B-2 (1-569) x CA985560 (1-924)

246 IleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGlu 265

31 ATCGAGTCCAACTGGCGCTGGGAGACACAGCTTCAAGATTCTACGCGAGTGAA 90

266 ThrSerLysGlyValTrpCysLeuGlnTrpAspAspGlnLysIleValSerGlyLeuArg 285

91 ACAAGCAAGGGGTATCTGCTCGATGATCAGAGATAGTAAGTGGACTCAGA 150

286 AspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGly 305

151 GATACACCATTAAGATCTGGGATTAAGAACTTTGGAGTGCAGGAGTGCTGATGGGT 210

306 HisThrGlySerValLeuCysLeuGlnTrpAspGluArgValIleLeuThrGlySerSer 325

211 CACACTGGGTCTCTCTGCTCGATGATGAGAGATATCATTTACTGCTCTTCA 270

326 AspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHis 345

271 GACTCTACCGTCCGGGTGGGAGCGTGAACACAGAGAAATGTGACACGCTGATTCAC 330

346 HisCysGluAlaValLeuHisLeuArgPheAsnGlyMetMetValThrCysSerLys 365

331 CACTGTGAGGTGCTGCTGCACTTGAGGTTTAATATGCGCATGATGGTCACCTCTCAA 390

366 AspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgVal 385

391 GATGCTTCATTCGAGTTTGGGACATGCGCTCTCGACCGCATATCATATTACGAGATT 450

386 LeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleValSer 405

451 CTGGTAGCCACCGAGTCGCGTAATGTGTGGACTTTGATGACAAAGTATAGTCTCT 510

406 AlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThr 425

511 GCATCTGGTATCGAACAATTAAGGTTTGGAAACACCATGATGATGATTTGTGCGACA 570

426 LeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTrpArgAspArgLeuValSer 445

571 TTGAACGCCCAACAGCGTGGTATTGTCATGCTTACAGATCGAGATCGGCTTGTGGT 630

446 GlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgVal 465

631 GGTCTTCTGACACACAAATCAGATTGGGGATTAATGATGCGGTGATGTTTGGGGTT 690

466 LeuGluGlyHisGluGluLeuValArgCysIleArgPheAsnLysArgIleValSer 485

691 CTGAGGAGCATGAGAACTGTCGCGTGCATCCGCTTTGATACAGAGAAATAGTCAGT 750

486 GlyAlaTrpAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAla 505

751 GGAGCATATGACGGGAAATTAAGTGTGGGACTTGTGCTGCTTTGGACCCCGTGGC 810

506 -ProAla-GlyThrLeuCysLeuArg-ThrLeu-ValGluHisSerGlyArgValPhe-- 523

811 GCTTGCAGGAGCCCTGTGTCTCCGGACTCTTTTGGAGCATTCGGGCAGAGATTTTC 870

524 -ArgLeuGln-PheAspGluPhe 530

871 CGCCTTGACGTTTTCAGGAGTTT 893

RESULT 7

AL898993/c

LOCUS

DEFINITION

AL898993 XGC-egg silurana tropicalis cDNA clone TEG9019e21 3', mRNA

sequence.

AL898993

ACCESSION

AL898993.1 GI:22951385

VERSION

KEYWORDS

SOURCE

Silurana tropicalis (western clawed frog)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 895)

Taylor, R., Ashurst, J.L., Croning, M.D.R., Zorn, A.M. and Rogers, J.

Sanger Xenopus tropicalis EST project 2002

Unpublished

COMMENT

Contact: Taylor R

Sanger Centre

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCEID: TEG9019e21.q1k17

Sequencing primer: T7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

Location/Qualifiers

Source

1-895

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TEG9019e21"

/dev_stage="egg"

/lab_host="Escherichia coli XL1-blue"

/clone_lib="XGC-egg"

/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA

was oligo dT primed from Sug of poly A+ RNA from egg.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end"

BASE COUNT 232 a 232 c 199 g 232 t

ORIGIN

Alignment Scores:

Pred. No.: 9.46e-143 Length: 895

Score: 1390.00 Matches: 264

Percent Similarity: 98.15% Conservative: 1

Best Local Similarity: 97.78% Mismatches: 5

Query Match: 45.81% Indels: 0

DB: 9 Gaps: 0

US-09-601-168B-2 (1-569) x AL898993 (1-895)

OY 300 LysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTrpAspGluArgVal 319

894 AAACGAGTGTGATGGGCCACATGGGTGAGTCTCTGCTGCAGTATGATGAGAGAGTA 835

OY 320 IleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMet 339

834 ATCAATACCGGCTCTTCTGACTCTACCGTCGGGTTGGGACGTAAACACAGAGAAATG 775
 340 LeuAsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAnGlyMet 359
 774 TTGAACACTCTGATTACCACTGTGAGGCTGTGCTACATTGAGGTTTAAACAATGGTATG 715
 360 MetValThrCysSerLeuAspArgSerIleAlaValTrpAspMetAlaSerProThrAsp 379
 714 ATGGTCACCTGCTCCAAAGATCGTTCCATTGCAGTGTGGACATGGCGCTTCGCCACAGAT 655
 380 IleThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAsp 399
 654 ATCACTATACGAAGAGTTCTAGTAGCCACCGAGCTGCCGTAATGTGGACTTTGAT 595
 400 AspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThr 419
 594 GACAAGTACATAGTCTCGCTCTGGTGTGATCGAACATAAAGGTTTCGAACACAGTACA 535
 420 CysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArg 439
 534 TGTGAATTGTGGCGACATTGAATGCCACAGCGTGGTATTGCTTACAGTATCGA 475
 440 AspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCys 459
 474 CATAGGCTTGTGGTGAAGTGTCTCTCATACCAATCAGATTGTGGACATTGAATGC 415
 460 GlyAlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAsp 479
 414 GGAGCATGTTTGGGGTCTCGMAGGACATGAAGACTGGTTGATCGATTCGCTTTGAT 355
 480 AsnLysArgIleValSerGlyValTyrAspGlyLysIleLysValTrpAspLeuValAla 499
 354 AACACAGATAGTCAGTGGAGCATATACCGGGAATTTAAAGTGTGGGACCTTGTGCT 295
 500 AlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSer 519
 294 GCTTTAGACCCCGGACCTCGACGGACCTGTGTCTCCGACTCTGTGGAGCATCG 235
 520 GlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAsp 539
 234 GGCAGAGTTTCCGCTCGCTGATTTGATCAGTTCAGATAGTAAGCAGCTCTCAGCAGC 175
 540 ThrIleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSer 559
 174 ACTATCCCATCTCGGATTTCTCAACACCCAGCTGCACATGCAGAGTCAACGCGTTC 115
 560 ProSerArgThrTyrThrTyrIleSerArg 569
 114 CTTCTAGACATACACGTACATCTCCAGA 85

8 2
 CA980902 966 bp mRNA linear EST 27-FEB-2003
 ION AGENCOURT 11279402 Wellcome CRC psk egg Xenopus laevis cdna clone
 ON IMAGE:6869574 5', mRNA sequence.

CA980902.1 GI:27513556

EST.

Xenopus laevis (African clawed frog)

ISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.

1 (bases 1 to 966)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

RS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

AL Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: N. Garrett, P. LeMaire, A.M. Zorn, and J.B.

Gordon (Wellcome/CRC Institute)

cdna Library Preparation: N. Garrett, P. LeMaire, A.M. Zorn, and

J.B. Gordon (Wellcome/CRC Institute)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: XGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Place: LUM4498 row: j column: 05
 High quality sequence start: 15
 High quality sequence stop: 619.
 Location/Qualifiers

FEATURES

source

1. 966

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:6869574"

/tissue_type="egg"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Wellcome CRC psk egg"

/note="Vector: pBluescript SK-; Site 1: NotI; Site 2:

EcoRI; cDNAs were oligo-dT primed and directionally

cloned. Library was constructed by N. Garrett, P. LeMaire,

A.M. Zorn, and J.B. Gordon (Wellcome/CRC Institute).

Note: This is a Xenopus Gene Collection (XGC) library."

BASE COUNT 256 a 202 c 269 g 239 t

ORIGIN

Alignment Scores:

Pred. No.: 1.32e-138 Length: 966

Score: 1353.00 Matches: 275

Percent Similarity: 91.80% Conservative: 5

Best local Similarity: 90.16% Mismatches: 19

Query Match: 44.59% Indels: 8

DB: 14 Gaps: 3

US-09-601-168b-2 (1-569) x CA980902 (1-966)

QY 228 ProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGlu 247
 Db 10 CCAAAATTCATTCTCCAGCGCTTTACCCAAAATTTATTCAGACATAGAGCAATCGAG 69
 QY 248 SerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSer 267
 Db 70 TCCAACTGGCGCTCTGGGAGACACAGCTTACAAGAAATTCACCTGCCGGAGTGAACAAGC 129
 QY 268 LysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsn 287
 Db 130 AAAGGGGTATACCTCTGCAGTACGATGATCAGAGATAGTAAGTGGACTCAGAGATAAC 189
 QY 288 ThrIleLysIleTrpAspLysAsnThrIleGluCysLysArgIleLeuThrGlyHisThr 307
 Db 190 ACCATTAAAGATCTGGGATAAGAATCTTTGGAGTCAAGCGAGTGTGATGGGTCACT 249
 QY 308 GlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSer 327
 Db 250 GGGTCAGTTCTCTCTGCAATATGATGAGAGAGTAATCATTAATGCTCTTCAGACTCT 309
 QY 328 ThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrIleLeuHisHisCys 347
 Db 310 ACCGTCCGGGTGTGGGACGTGAACACAGGAGAAATGTTGAACACGCTGATTCACCACCTGT 369
 QY 348 GluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArg 367
 Db 370 GAGGCTGTGGCTGCACCTTGAGGTTTAATAATGTCATGATGTCACCTGCTCCAAAGATCGT 429
 QY 368 SerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuVal 387
 Db 430 TCCATTGCAGTTTGGGACATGGCTCTCGCCACCGATATCATCATACGAGAGTCTGGTA 489
 QY 388 GlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleValSerAlaSer 407
 Db 490 GGCCACCCGAGCTGCCGTAATGTGGTGGACTTTTGATGACAAGTATATAGTCTCTGCACT 549
 QY 408 GlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsn 427

```

|||||
550 GGTGATCGAACATAAAGTTTGGACACCATGATGATGATTTGGCGGACATTGAAC 609
|||||
428 GlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGlySer 447
|||||
610 GGCCACAAAGCGTGGTATTGCATGCTTACAGTATCGAGATCGGCTTGTGGTGAAGTGTCT 669
|||||
448 SerAspAsnThrIleArgLeuTyrAspIleGluCysGlyAlaCysLeuArgValLeuGlu 467
|||||
670 TGTGACAC-ACAAATCAGATTGGGATATTGATGCGGTGATGTTTCCGGGGTCTGGAA 728
|||||
468 GlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyAla 487
|||||
729 GGACATGAAGAACTGGTGGTGCATCCGCTTTGATAC-AAGAGAATAGTCAGTGGAGCA 787
|||||
488 TyrAsp-GlyLysIleLysValTyrAspLeu-ValAlaAlaLeuAspProArgAla---P 506
|||||
788 TATGACGGGAAAAAATAAGTGTGGGAACCTTTGTCGCTCTTTTGGGCCCCCGGGCGGC 847
|||||
506 roAlaGlyThr-----LeuCysLeuArgThrLeuValGluHis---SerGlyArgValP 523
|||||
848 CTCAGGGGACCCCTGGTCTCCCGGAATCCTTGGTGAACATTTCCGGGCGAGAGGTTT 907
|||||
523 heArgLeuGln 526
|||||
908 TCCCCCTGGAA 918

```

```

9
26
CB244526 776 bp mRNA linear EST 12-FEB-2003
UI-M-FYO-cdgr-k-22-0-UI.r1 NIH_BMAP_FYO Mus musculus CDNA clone
IMAGE: 6833135 5', mRNA sequence.
CB244526
EST.
CB244526.1 GI:28366170
Mus musculus (house mouse)
Mus musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 776)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..776
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6833135"
/tissue_type="whole brain"
/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FYO"
(note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated

```

with EcoR I adaptor, digested with NotI and then cloned directionally into pfx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGCAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 201 a 185 c 212 g 175 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 2e-138 Length: 776
Score: 1350.00 Matches: 252
Percent Similarity: 98.06% Conservative: 1
Best Local Similarity: 97.67% Mismatches: 5
Query Match: 44.50% Indels: 0
DB: 14 Gaps: 0

US-09-601-168B-2 (1-569) x CB244526 (1-776)

QY	152	GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla	171
Db	3	GGTCTGGACCAACATCGCTGAGAACATTTCTGTACATCTTGACCGCCAGTCACGTGTGTCT	62
QY	172	AlaGluLeuValCysLysGluTyrArgValThrSerAspGlyMetLeuTyrLys	191
Db	63	GCTGAGCTCGTGTCAAAGGAATGTCACCGGTGACGTCCGACGGCATGCTGTGGAAAAAG	122
QY	192	LeuIleGluArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGluArgGly	211
Db	123	CTCATCGAGAGGATGGTCAGACGGACTCTCTGTGGCGAGCGCTGCGAGCGCAGAGGC	182
QY	212	TyrGlyGlnTyrLeuPheLysAsnLysProAspGlyAsnAlaProProAsnSerPhe	231
Db	183	TGGGACAGACTATTATTCAAAAACAAACCTCTGTGTGAGAACGCTCTCTCCCACTCTTT	242
QY	232	TyrArgAlaLeuTyrProLysIleIleGluAspIleGluThrIleGluSerAsnTyrArg	251
Db	243	TATAGAGCGCTTTATCTCTAAATCATACAAGACATTTGAGACAATAGAGTCCAAATTGAGA	302
QY	252	CysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyr	271
Db	303	TGTGGCGCACATAGTTTACAGAGATCCACTGCGGAGTGAACAAAGTAAGGGGTTTAC	362
QY	272	CysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIle	291
Db	363	TGTTTACAGTACGACGACCAAGATAGTCAGCGGCTTCGAGACAACACCATCAAGATC	422
QY	292	TyrAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeu	311
Db	423	TGGGATAAAGCACACTGGATGCAAGCGGATTTCTCAGGGGCCACACGGGCTCCGTCCTG	482
QY	312	CysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThrValArgVal	331
Db	483	TGTTCTGAGTACGATGAGAGGGTATCATCACAGGCTCTCTCAGACTCCACCGTCAGAGTG	542
QY	332	TyrAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaValLeu	351
Db	543	TGGGATGTAAATGCGGTGAGATGCTAAACACATTTGATTCACCACCTGTGAAGCGCTCTG	602
QY	352	HisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaVal	371
Db	603	CACCTGGCTTCANTAAATGTCATGATGGTGACCTGTTCANAGACCGCTTCATCGCTGTG	662
QY	372	TyrAspMetAlaSerProThrAspIleThrLeuArgValLeuValGlyHisArgAla	391
Db	663	TGGGATATGGCTTCCCACTGACATCACCTCCANAGAGGGTGTGTGGACACCGAGCT	722
QY	392	AlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSerGlyAsp	409
Db	723	CGGTCATATGTTGTAGACTTTGATGACAAAGTATCATGTTTCTGCTCTGGAGAT	776

Seq primer: pYX-5.
Location/Qualifiers
1..776
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6833135"
/tissue_type="whole brain"
/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FYO"
(note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated

```

BM944304          810 bp      mRNA      linear      EST 14-MAR-2002
UI-M-EHOp-bvg-m-18-0-UI.r1 NIH_BMAP_EHOp Mus musculus cDNA clone
IMAGE:5695673 5', mRNA sequence.

BM944304          810 bp      mRNA      linear      EST 14-MAR-2002
BM944304.1      GI:19427889
EST.

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
1 (bases 1 to 810)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. 810
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5695673"
/tissue_type="whole brain"
/dev_host="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EHOp"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is CAGCCAGGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

JNT      189 a      184 c      221 g      216 t

t Scores:
3.56e-138      Length:      810
1348.00      Matches:      252
Similarity:      97.78%      Conservative:      12
al Similarity:      93.33%      Mismatches:      6
tch:      44.43%      Indels:      0
12      Gaps:      0

1-168B-2 (1-569) x BM944304 (1-810)

276 AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsn 295
1 GATGATCACAAATATATCACTGGCTCTCGGACCACTCTATCAAGTCTGGGATTAAGC 60

296 ThrLeuGluCysLysArgLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyr 315
:::

```

Db	61	AGCTTGGAAATGTTTGAAGATGCTAAACGGGCCACACAGGCTCTGTCTCTCTGCTCCAGATAT	130
QY	316	AspGluArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsn	335
Db	121	GATGAGCGAGTCATCTGTAAGTCTTCCAGACTCCACGGTGAGAGTCTGGGATGTGAAC	180
QY	336	ThrGlyGluMetLeuAsnThrIleuIleHisHisCysGluAlaValLeuHisLeuArgPhe	355
Db	181	ACTGGTGAGGTGCTCAACACACTCATCCACCAATGAAGCGGTACTGCACTTACCGCTTC	240
QY	356	AsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAla	375
Db	241	AGCAATGCAGCTGATGTGACTTGTTCAGGACCGTTCCTCAATGCGGTGGACATGGCT	300
QY	376	SerProThrAspIleThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnVal	395
Db	301	TCTGCCACCGATATCACTTTAGCCGGTGTCTGTTGGCCACCGTGTCTGCTCAATGTA	360
QY	396	ValAspPheAspAspLysVrIleValSerAlaSerGlyAspArgThrIleLysValTrp	415
Db	361	GTAGACTTTGATGATAAATACATCGTGTCTGCTTCAGAGACAGGACCATTAAGTGTGG	420
QY	416	AsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCys	435
Db	421	AGCAGCAGCATGTGAGTTGTTCGCACCTCTGAAATGGGCAACAGCGGATCGCGTGT	480
QY	436	LeuGlnTrpArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrp	455
Db	481	CTGAGTACCGCCAGCCGGCTTGTGTGTAGTGATCATCAGATATATACATCCGGTTATGG	540
QY	456	AspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCys	475
Db	541	CATATTCAATGTGTCCTGTTTAAAGATGCTTAGAGCGCCACGAAGAATTTGGTCCGGTGC	600
QY	476	IleArgPheAspAsnLysArgIleValSerGlyAlaTrpAspGlyLysIleLysValTrp	495
Db	601	ATCCGTTTGTATACAGAGAGATGTCACTGGCGCTATGATGGAGATTAAGTCTGG	660
QY	496	AspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeu	515
Db	661	CACTTGCAGGCTGCTCTTGACCCCTCGGCCCCAGCAACACATGTGTCTGCGCACCTTG	720
QY	516	ValGluHisSerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSer	535
Db	721	GTGGAACACACTCTGCAGCTGTGTTTCGGCTGCGATTTTGATGAGTTTCAGATCATCAGCAGC	780
QY	536	SerHisAspAspThrIleLeuIleTrpAsp	545
Db	781	TCCCATGATGACACTATTGTATTGGGAT	810
RESULT 11	BQ444188	778 bp mRNA linear EST 29-MAY-2002	
LOCUS	UI-M-BX0-bxj-1-06-0-UI.r1 NIH_BMAP_EXO	Mus musculus cDNA clone	
DEFINITION	IMAGE:5709077 5', mRNA sequence.		
ACCESSION	BQ444188		
VERSION	BQ444188.1	GI:21247300	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	EuKaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
AUTHORS	NIH-MGC	http://mgi.nci.nih.gov/	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-x@mail.nih.gov Tissue Procurement: Dr. James Lin, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be		

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>
(This clone was contributed by the Brain Molecular Anatomy Project (BMAP))

Seq primer: pYX-5.

Location/Qualifiers

1..778

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:570907"

/tissue_type="whole brain"

/dev_stage="embryo 15.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NH MAP EXO"

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 5:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction. ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

223 a 181 c 204 g 163 t 1 others

JUNT

nt Scores:

O.: 1.96e-136 Length: 778
1332.00 Matches: 253
Similarity: 98.07% Conservative: 1
cal Similarity: 97.68% Mismatches: 5
atc: 43.90% Indels: 1
13 Gaps: 0

01-168B-2 (1-569) x BQ444188 (1-778)

63 AlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaSerMet 82

1 GCTATGAGACTGAAATGTGTGGCCAAAGCCAACTTGCATGTCATTCAGCATG 60

83 IleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysVal 102

61 ATTTGCCAAGCAGCGGAACTCTCAGCAAGCTATGAGAGGAAAGGAGCTGTGTGTC 120

103 LysTyrPheGluGlnTyrSerGluSerAspGlnValGluPheValGluHisLeuIleSer 122

121 AAGTATTGTGACGTGTCTGAGTCTGATCAAGTGGATTTGTAGAACCTTATATCC 180

123 GlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArg 142

181 CAAATGTGTCACTACCACTGGGCACATCACTCACTTAAACCTATGCTGCAGAGG 240

143 AspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSer 162

241 GATTTTCAATACTGCATCCAGCAGCGGGCTGGACACATCGCTGTGAGAACATTTGTCA 300

163 TyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyrArgVal 182

301 TACTTGGACGCCAAGTCACTGTGTCTCTGAGCTGTGTGTCAAGGAATGTCACCGGTG 360

183 ThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeu 202

361 ACGTCGGACGGCATCTGTGGAAAAGCTCATCGAGAGGATGGTCAGACGGACTCTCTG 420

203 TrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProPro 222

Db 421 TGGCGAGGCTTGGCAGAGCGCAGAGGCTGGGACAGTACTTATCAAAAACCACTCCT 480

Qy 223 AspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAsp 242

Db 481 GATGAGAACGCTCTCTCCAACTCTTTTATAGAGCGCTTTATCTTAAATCATACAAGAC 540

Qy 243 IleGluThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIleHisCys 262

Db 541 ATTGAGACAATAGATCCATTTGGAGATGTGGCGGACATAGTTTACAGAGATCCACTGC 600

Qy 263 ArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSer 282

Db 601 CGAGTGAACAAGTAAGGGGTTTACTGTTTACAGTACGAGCAGCAGAGATAGTCAGC 660

Qy 283 GlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLysArgIle 302

Db 661 GGCCTTCGAGACAC-ACCATCAAGATCTGGANTAAAGCACACTGGAAATGCAAGCGGATT 719

Qy 303 LeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIle 321

Db 720 CTCACGGGCCACACAGCGGCTCCGTCCTGTGTCTGCTGATGATGATGATGATC 776

RESULT 12

CA971789

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA971789 1002 bp mRNA linear EST 27-FEB-2003
AGENCOURT 10768214 Wellcome CRC PSK egg Xenopus laevis cDNA clone
IMAGE:6323303 5', mRNA sequence.

CA971789
CA971789.1 GI:27504443
EST.

Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.

1 (bases 1 to 1002)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
Gordon (Wellcome/CRC Institute)
cDNA Library Preparation: N. Garrett, P. LeMaire, A.M. Zorn, and
J.B. Gordon (Wellcome/CRC Institute)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: XGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LAM3766 row: d column: 24
High quality sequence start: 98
High quality sequence stop: 710.

FEATURES

source

1..1002
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6323303"
/tissue_type="egg"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Wellcome CRC PSK egg"
/note="Vector: pBluescript SK-; Site 1: NotI; Site 2:
EcoRI; cDNAs were oligo-dT primed and directionally
cloned. Library was constructed by N. Garrett, P. LeMaire,
A.M. Zorn, and J.B. Gordon (Wellcome/CRC Institute).
Note: This is a Xenopus Gene Collection (XGC) library."
296 a 207 c 259 g 239 t 1 others

BASE COUNT

ORIGIN

Alignment Scores:

IMAGE:6802203 5', mRNA sequence.
CA975032
CA975032.1 GI:27507686
EST.
Danio rerio (zebrafish)
Danio rerio
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
DNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUAM4323 row: c column: 02
High quality sequence start: 24
High quality sequence stop: 652.
Location/Qualifiers
1. 935
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6802203"
/tissue_type="embryo"
/lab_host="DH10B (TI-resistant)"
/clone_lib="NCI CGAP ZEMB2"
/notes="Vector: PCMV-SPORT6 cdb; Site: 1; EcoRV; Site: 2;
Noti; Cloned unidirectionally. Primer: Oligo 47. Average
insert size 2 kb. Constructed by J. Wang (Research
Genetics, Invitrogen Corp) from tissue donated by L. Zon
(Harvard University). Note: this is a NCI CGAP Library."
223 a 214 c 265 g 232 t 1 others

BASE COUNT 223 a 214 c 265 g 232 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 935
Score: 1320.50
Matches: 266
Percent Similarity: 92.18%
Best Local Similarity: 86.64%
Query Match: 43.52%
DB: 14
Gaps: 1

US-09-601-168B-2 (1-569) x CA975032 (1-935)

QY 247 GluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgHisCysArgSerGluThr 266
:::
Db 25 GATNCAACCTCGAGGTGGTGGACATACCTCGAGAGGATTCAGTGTAGGTGGAGAAC 84
QY 267 SerLysGlyValTyrCysLeuGlnTyrAspAspClnLysIleValSerGlyLeuArgAsp 286
Db 85 AGCAAAAGGTGTTACTGCTCCAGTACGATGACGAGAAATCATCATCGGTCTACGAGAC 144
QY 287 AsnThrLysIleTyrAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHis 306
:::
Db 145 AATTCCATCAAGATATGGATAAGACACATCTCGAGGTGTTGAAGATCTTGACGGGAC 204
QY 307 ThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleLeuThrGlySerSerAsp 326
Db 205 ACGGGCTCTGTGCTGTCTGCAATACGATGAGAGAGTCAATGTGTACCGGTCTCTCTGAT 264
QY 327 SerThrValArgValTyrAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisHis 346
Db 265 TCCACTGTGAGAGTGTGGATGTGAATTCGGGGGAGGTGTTTGAACATCTGTATCCATCAC 324
QY 347 CysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAsp 366

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1002
Length: 1002
Matches: 258
Conservative: 10
Mismatch: 14
Indels: 8
Gaps: 1

601-168B-2 (1-569) x CA971789 (1-1002)

93 SerTyrGluLysGluLysGluLysValLysTyrPheGluGlnTrpSerGluSerAsp 112
:::
19 AATTACGAGAGAGAGAGAGCTATGCTCAAGTATTTGAGCAGTGGTCCGAGTGGAT 78
113 GlnValGluPreValGluHisIleSerGlnMetCysHisTyrGlnHisGlyHisIle 132
:::
79 CAAGTAGAGTTGTTGAACACCTGATCTTCCGAATGTCACATATCAGTATGACATATA 138
133 AsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGly 152
:::
139 AACACTTACCTTAAGCCCAATGTTCAAAAGGATTTTCATCCCGCACTACCGAGTCGCGGA 198
153 LeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAla 172
:::
199 CTAGATCATAGCAAAACATCTTTTACCTGACGCAAGTCAATGTGTTTCGCA 258
173 GluLeuValCysLysGluTyrTyrArgValThrSerAspGlyMetLeuTyrLysLysLeu 192
259 GAACCTGGTATGTAAAGAGTGTGTACAGAGTGACCTCTGATGGAATGCTCTGGAAGAGCTC 318
193 IleGluArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGluArgArgGlyTyr 212
319 ATAGAGAGAGTGTGTGAGAGAGTCTCTTTTGGAGGGGACTTGGCAGAGAAGAGAGATGG 378
213 GlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAlaProProAsnSerPheTyr 232
379 GTCTCAATATCTATTAAACCAAACTCCAGATGGGAAACCCACCAATTCCTCTTAC 438
233 ArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCys 252
439 AGAGCACATTTACCCAAAAATTATTCAACATAGACATAGACATAGAGTCCAACTGCGCTGT 498
253 GlyArgHisSerLeuGluArgIleHisCysArgSerGluThrSerLysGlyValTyrCys 272
499 GGGAGACACAGCTTCAAAAGAAATTCACCTGCGGAGTGAACAACGAAAGGGGAACTGT 558
273 LeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyr 292
559 CTGCGAGTACGATGATCAGAGATAGTAAGTGGACTCAGAGATAACACATTAAGATCTGG 618
293 AspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCys 312
619 GATAAGAACTTTTGGAGTGAAGGAGTGTGATGGCCACACTGGGTCTAGTCTCTGT 678
313 LeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThrValArgValTyr 332
679 CTGCGATGATGAGAGAGTATCATTTACCGGCTCTTACAGCTCTACTGCTCCGGGTTG 738
332 pAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisHisCysGluAlaVal 351
739 GGACGTGAACACAGGAAGAAATGTTGAACACACATTTGATTCACCACTGTGGAGGCGCTGTC 798
351 euHisLeuArg--PheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIle 369
799 TGCATTTGGAGGTTTAAACAAATGGATGATGGGTGACCTTGGCTCCCAAAAGATCTTT 858
369 eAlaValTyrAspMet 374
:::
859 TTCCATTTGGCAGTTG 874

CA975032 935 bp mRNA linear EST 06-JAN-2003
ON AGENCOURT 11076265 NCI CGAP ZEMB2 Danio rerio cdna clone

```

325 AACGAGCGCGTCCACCTGCGTTCGTAAAGGGCTGATGTCACCTGCTCAAGAGAT 384
|||||
367 ArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgValLeu 386
|||||
385 CGCTCCATTGAGTGTGGGACATGGCTCTCCACACACATAAGCGCTCCCGGGTACTC 444
|||||
387 ValGlyHisArgAlaValAsnValValAspPheAspIleValSerAla 406
|||||
445 GTTGACACACAGCTGACGCTCAATGCTGATTTGATGACAGTACATGCTCTGCGC 504
|||||
407 SerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeu 426
|||||
505 TCTGGGATCGGACTATAAGGTATGGAGTACAAAGCACCTGCGAGTTTGTTCGACACTA 564
|||||
427 AsnGlyHisLysArgGlyIleAlaCysLeuGlnTrpArgAspArgLeuValSerGly 446
|||||
565 AATGCTCATAAGCGAAGAAATTGCTGTGCTGAGTACAGACCGGACTTGTGTCAGCGGC 624
|||||
447 SerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeu 466
|||||
625 TCTTCAGACACACCATCAGCTGCTGGGATATTGAGTGGGAGCATGTTTGGGGTTTG 684
|||||
467 GluGlyHisGluGluLeuValArgCysIleArgPheAspIleValSer-Gl 486
|||||
685 GGAGGACACGAGGAGCTTGTTCGCTGATTTCGCTTTGACAAACAAACGATCGTGGTGG 744
|||||
486 YAlaTrpAspGlyLysIleLysValTrpAspLeuValAlaAlaLeu-AspProArgAlaP 506
|||||
745 AGCGTATGACGGTAAATTTAAGTATGGGACCTACAGCTGCTGCTTTCACCGGCTC 804
|||||
506 roAlaGlyThrLeuCysLeuArgThrIleVal-GluHisSerGly--ArgValPheArgL 525
|||||
805 CAGCCAGACACTCTGTCTACGCCCTTAGTGGGAACATTCAGGGCCGAAAGTGTTCAGGG 864
|||||
525 euGln-PheAspGlu-PheGlnIleVal-SerSerSerHisAspAspThrIle--LeuI 543
|||||
865 TTCAGTTTGGAGGTTTCGAAATCATCCAAAGTTCCTGATGATGAACCTTATCCCTCA 924
|||||
543 leTtp 544
|||||
925 TCTGG 929
|||||

14
1 CA327271 850 bp mRNA linear EST 27-NOV-2002
ION UI-M-FY0-ccy-f-17-0-UI.r1 NIH_BMAP_PYO Mus musculus cDNA clone
ON IMAGE: 6826098 5', mRNA sequence.
CA327271
S EST.
S Mus musculus (house mouse)
ISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CE 1 (bases 1 to 850)
RS NIH-MGC http://mgi.nci.nih.gov/.
AL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5 Location/Qualifiers
1. .850

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6826098"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"
/lab_hosts="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_PYO"
/notes="Organ: Brain; Vector: pYX-Asc; Site1: EcoR I; Site2: Not I; The library was constructed according to Ronald, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT 221 a 190 c 222 g 212 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 134 Length: 850
Score: 1310.00 Matches: 247
Percent Similarity: 92.63% Conservative: 17
Best local Similarity: 86.67% Mismatches: 19
Query Match: 43.18% Indels: 3
DB: 14 Gaps: 1

US-09-601-168B-2 (1-569) x CA327271 (1-850)
QY 178 GluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLeuIleGluArgMetVal 197
DB 3 GAATGGCAGCGAGTGATCTCAGAAGGGATGCTTTGGAAGAGCTGATTGAGAGATGGTG 62
QY 198 ArgThrAspSerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPhe 217
DB 63 CGCACCGACCTCTCTGGAAGGACTCTCAGAAGAGAGGGCTGGGATCAGTACCTGTTT 122
QY 218 LysAsnLysProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrPro 237
DB 123 AAAAACAACACCTACAGATGGC-----CCTCCCAACTCATTTTATAGATCAATTATACCA 176
QY 238 LysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeu 257
DB 177 AAGATTATCCAGACATAGACCATAGATCCAACTGCGGCTGTGGACACCAACTTG 236
QY 258 GlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAsp 277
DB 237 CAGAGGATCCAGTCCGCTCTGAAAATAGTAAGGGTGTCTACTGTTTGAATATGATGAT 296
QY 278 GlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeu 297
DB 297 NACAAATATATCATGGCCCTCCGGGCAACTCATCAAGATCTGGGATAAAGCAGCTTG 356
QY 298 GluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu 317
DB 357 GAATGTTTGAAGTGCTTAACGGGCCACACAGGCTCTCTCTCTGCTCCAGTATGATGAG 416
QY 318 ArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGly 337
DB 417 CGAGTCATTTGTAATGTTTCTTCAGACTCCACCGGTGAGAGTCTGGGATGTGAACACTGGT 476
QY 338 GluMetLeuAsnThrIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsn 357
DB 477 GAGGTGCTCAACACACTCATCCACCAATGAAGCGGCTACTGCACCTTACGCTTCAGCAAT 536

```


358 GlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerPro 377
 537 GCAGTGTGTGACCTGTTCACAGACCGTTCCTATTCCTGCTGGACATCGCTTCTGCC 596
 378 ThrAspIleThrLeuArgAcValLeuValGlyHisArgAlaAlaValAsnValAsp 397
 597 ACCGATATCATCTTACGCCCGTGTCTGTGGCCACCGTGTCTGTCAATGTAGTANAC 656
 398 PheAspAspLysTyIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThr 417
 657 TTGTATGATNAATACATCGTCTGCTTCAGGACAGGACCAATTAAGTGTGGAGCAG 716
 418 SerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGln 437
 717 AGCATGTGTGAGTTTGTCCGCACTCTGAATGGGCACAGCGAGCATCGCTGTGTGCAN 776
 438 TyrArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIle 457
 777 GTACGCGACCGCTGTGTGTGTAGTGATCATCNAGATATACATC-CGGTTATGGGATATT 835

458 GluCysGlyAlaCys 462

836 TGATGTGTGCTGT 850

15

36 CB520936 744 bp mRNA linear EST 28-MAR-2003
 UI-M-GH0-cel-p-18-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
 IMAGE: 6841315 5', mRNA sequence.

ON CB520936

35 CB520936.1 GI:29354291

35 EST.

35 Mus musculus (house mouse)

35 Mus musculus

35 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

35 NIH-MGC <http://mgi.nci.nih.gov/>.

35 National Institutes of Health, Mammalian Gene Collection (MGC)

35 Unpublished

35 Contact: Robert Strausberg, Ph.D.

35 Email: cgabbs-remail.nih.gov

35 Tissue Procurement: Dr. Jim Lin, University of Iowa

35 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

35 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

35 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

35 Clone Distribution: Distribution information can be found at

35 <http://genome.uiowa.edu/distribution/mousefl.html>

35 This clone was contributed by the Brain Molecular Anatomy Project

35 (BMAP)

35 Seq primer: pYX-5.

35 Location/Qualifiers

35 1..744

35 /organism="Mus musculus"

35 /mol_type="mRNA"

35 /strain="C57BL/6"

35 /db_xref="taxon:10090"

35 /clone="IMAGE: 6841315"

35 /tissue_type="whole brain"

35 /dev_stage="1, 5, and 15 days newborn"

35 /lab_host="DH10B (T1 phage resistant)"

35 /clone_lib="NIH_BMAP_GH0"

35 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

35 Site 2: Not I; The library was constructed according

35 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

35 1996. Denatured RNA was size fractionated on a 1% agarose

35 gel. First strand cDNA synthesis was primed with oligo-dT

35 Primer containing a Not I site. Double strand cDNA was

35 size selected according to mRNA size fraction, ligated

35 with EcoR I adaptor, digested with NotI and then cloned

35 directionally into pYX-Asc vector. The library tag

35 sequence located between the Not I site and the polyA tail

35 is CGAACTGAAT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

BASE COUNT 197 a 184 c 202 g 160 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.37e-133 Length: 744

Score: 1306.00 Matches: 243

Percent Similarity: 98.79% Conservative: 1

Best Local Similarity: 98.38% Mismatches: 3

Query Match: 43.05% Indels: 0

DB: 14 Gaps: 0

US-09-601-168B-2 (1-569) x CB520936 (1-744)

QY 145 IleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeu 164
 DB 2 ATAACCTGCACCTGCCAGCACGGGTCTGGACCACATCGTGAGAACATTTCTGTCATCTTG 61
 QY 165 AspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSer 184
 DB 62 GACGCCAAGTCACTGTGTCTGTGAGCTCGTGTCAAGGAATGGTACCGCGTGACGCTCG 121
 QY 185 AspGlyMetLeuTrpLysLysLeuIleGluAtrMetValArgThrAspSerLeuTrpArg 204
 DB 122 GACGCGCATCTGTGTGAAAAAGCTCATCGAGAGATGGTCAGACGACTCTCTGTGGCGA 181
 QY 205 GlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGly 224
 DB 182 GGCCTGGCAGACGCGACAGGCTGGGACAGTACTTATTCAAAAACAACCTCCTGATGAG 241
 QY 225 AsnIaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGlu 244
 DB 242 AACCTCTCTCCAACTCCTTTTATAGAGCGCTTTATCCTATAAATCATACAAGACATGAG 301
 QY 245 ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
 DB 302 ACATAGATCCCATTCGAGATGTGGCGACATAGTTTACAGAGATCCACTGCCGAGT 361
 QY 265 GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu 284
 DB 362 GAAACAAGTAAAGGGGTACTGTATTACAGTACGACGACGACAGAGATAGTCAGCGGCTT 421
 QY 285 ArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
 DB 422 CGACACACACCATCAAGATCTGGATAAAGCACACACTGGAGTGCAGAGCGGATCTCAG 481
 QY 305 GlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySer 324
 DB 482 GGCACACAGCGCTCCGTCTGTCTGTAGTACGATGAGAGGGTGTATCATCAGGCTCC 541
 QY 325 SerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIle 344
 DB 542 TCAGACTCCACCGTCAGAGTGTGGATGTAAATGCGAGTGTGATGCTAAACACATTTGATT 601
 QY 345 HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer 364
 DB 602 CACCACGTGTGAAGCGGTCTTCGCACCTCGCTTCAATAATGGCATGATGTGTGACCTGTTC 661
 QY 365 LysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArg 384
 DB 662 AAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCAACTGACATCACCTCCANGAGG 721
 QY 385 ValLeuValGlyHisArgAla 391
 DB 722 GTGCTGTGGGACACCGAGCT 742

Search completed: October 22, 2003, 17:11:13

Job time : 2690 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ein - nucleic search, using frame_plus_p2n model

October 22, 2003, 13:58:17 ; Search time 385 Seconds
(without alignments)
3989.559 Million cell updates/sec

US-09-601-168b-2

score: 3034
e: 1 MPFAEAVLOEKALFNMSS.....PAAQABPPSPRTTYVISR 569

table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 2552756 seqs, 1349719017 residues

umber of hits satisfying chosen parameters: 5105512

DB seq length: 0

DB seq length: 2000000000

ccessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

line parameters:

2.1/USPTO.spool.p/US0960168/runat.22102003.100545.17767/app_query.fasta_1.711
neseq 19Jun03 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
r=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
s -DOCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15
XAL -OUTPWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
09601168 @CGN_1_1_396 @runat.22102003.100545.17767 -NCPU=6 -ICPU=3
-LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
100T=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

: N Geneseq 19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

ed. No. is the number of results predicted by chance to have a
ore greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3034	100.0	1707	21	AAA73132	Human beta-transdu
2	3034	100.0	2151	21	AAK86501	CDNA encoding a hu
3	3034	100.0	2151	21	AAA51229	Human beta-TrCP co
4	3034	100.0	2151	21	AAZ93350	Sequence encoding
5	3034	100.0	2151	21	AAZ93710	F-box protein hBe
6	3034	100.0	2151	22	AAC84610	Human Zfil protein
7	3034	100.0	2151	24	AAAL1041	CDNA of Human F-bo
8	3034	100.0	2151	24	ABSS1009	Human cdna encodin
9	3034	100.0	2419	21	AAZ29233	Human cell signall
10	3037	99.8	3220	22	AAH90079	Human bone marrow
11	3036	99.1	2285	22	AAK51715	Human polynucleoti
12	2997	98.8	1707	21	AAA73131	Mouse ubiquitin li
13	2997	98.8	2175	21	AAZ93714	F-box protein FWD1
14	2993	98.6	3622	22	AAH89966	Human bone marrow
15	2932.5	98.6	2366	22	AAK51717	Human polynucleoti
16	2989.5	98.5	3003	22	AAK52699	Human polynucleoti
17	2989.5	98.5	3003	22	AAK52700	Human polynucleoti
18	2989.5	98.5	3003	22	AAK52701	Human polynucleoti
19	2892	95.0	2207	22	AAK51716	Human polynucleoti
20	2384.5	78.6	2884	22	AAI61150	Human polynucleoti
21	2384.5	78.6	4230	21	AAA51228	Human E3 ubiquitin
22	2384.5	78.6	4344	22	AAK52260	Human polynucleoti
23	2384.5	78.6	4360	22	AAI59364	Human polynucleoti
24	2089.5	68.9	2387	23	ABL06081	Drosophila melanog
25	1685	55.5	9129	23	ABL06080	Drosophila melanog
26	1163	38.3	657	24	ABK86903	Human beta TrCP (b
27	1155.5	38.1	951	23	ABK43155	CDNA encoding huma
28	935	30.8	1561	22	AAAC4601	Beta-TRCP.N/SKP2.C
29	702.5	23.2	594	23	ABV58856	Human prostate exp
30	689	22.7	479	22	ABA44075	Human breast cell
31	689	22.7	479	22	ABA54530	Human foetal liver
32	689	22.7	479	22	ABA24312	Probe #2778 for ge
33	689	22.7	479	22	AAK02821	Human brain expres
34	689	22.7	479	22	AAK28263	Human bone marrow
35	689	22.7	479	22	AAI12828	Probe #2761 for ge
36	689	22.7	479	22	AAI34184	Probe #2870 used t
37	689	22.7	479	22	AAI02746	Probe #2737 used t
38	689	22.7	479	23	ABS27866	Human liver single
39	689	22.7	479	24	ABS02776	Human genome-deriv
40	661	21.8	4135	23	ABL20843	Drosophila melanog
41	661	21.8	4146	23	ABL20841	Drosophila melanog
42	661	21.8	4158	23	ABL26469	Drosophila melanog
43	644.5	21.2	1881	20	AAZ99712	Human 6myc-N-sel-1
44	644.5	21.2	1881	22	AAF24136	Construct 6myc-N-s
45	640	21.1	1799	22	AAH14890	Human cdna sequenc

ALIGNMENTS

```

RESULT 1
AAA73132
ID AAA73132 standard; cDNA; 1707 BP.
XX
AC AAA73132;
XX
DT 27-NOV-2000 (first entry)
XX
DE Human beta-transducin repeat containing protein (beta-TrCP) cDNA.
XX
KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;
KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
KW gene therapy; colon cancer; beta-transducin repeat containing protein;
KW beta-TrCP; ss.
XX
OS Homo sapiens.
XX

```

2000166542-A.

-JUN-2000.

-DEC-1998; 98JP-0343437.

-DEC-1998; 98JP-0343437..

AGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

I; 2000-485550/43.

PSDB; AAB12813.

box protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or beta-catenin - sclosure; Fig 17; 19pp; Japanese.

e present invention describes an F-box motif protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or beta-catenin and is constituted by Skp1 protein, Cull1 protein and a repeat motif and has the amino acid sequence of 45 residues (AAB12811) one of two 569 residue sequences (AAB12812, which is mouse ubiquitin ligase FWD1 protein) and (AAB12813, which is human beta-transducin repeat containing protein (beta-TrCP)). The F-box protein can be used for a gene therapy of colon cancer by being recombined to a virus vector. e present sequence encodes the human beta-TrCP protein from the present invention.

quence 1707 BP; 514 A; 349 C; 414 G; 430 T; 0 other;

nt Scores:

O.: 4.02e-303 Length: 1707
 3034.00 Matches: 569
 Similarity: 100.00% Conservative: 0
 cal Similarity: 100.00% Mismatches: 0
 atch: 100.00% Indels: 0
 21 Gaps: 0

01-168B-2 (1-569) x AAA73132 (1-1707)

1 MetAspProAlaGluAlaValLeuGlnGluLeuAlaLeuLysPheMetAsnSerSerGlu 20
 1 ATGACCCCGCGAGGCGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 60
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysLeuLeuProGluLysAsnSerLeu 40
 61 AGAGAACACTGTAATAATGCGGAACCCCTTAGGAGATAATACCAGAGAAGAAATCACTT 120
 41 ArgGlnThrTyrrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 121 AGACAGACATACAAACAGCTGTCAGACTCTGCTTAAACCAAGAAACAGATGTTAGCA 180
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 181 AGCACTGCTATGAGACTGAGATTGTGTGCCCAACAAACAACTTGCCANTGGCACTTCC 240
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLysGluLysGluLeu 100
 241 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAAAGAAAGAACTG 300
 101 CysValLysTyrrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
 301 TGTGTCAAAATACTTGGAGCAGTGTGTGAGCTCAGATCAAGTGGAAATTTGTGGAACATCTT 360
 121 IleSerGlnMetCysHisTyrrGlnHisGlyHisIleAsnSerTyrrLeuLysProMetLeu 140
 361 ATATCCAAATGTCTCATACCAACATGGGCACATAAACTCGATCTTAAACCTATGTTG 420
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 421 CAGAGATTTTCATAACTGTCTGCCAGCTCGGGGATGGATCATATCGTGAGAACATT 480

Qy 161 LeuSerTyrrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrr 180
 Db 481 CTGTCTATACCTGGATGCCAAATCACTATGTGTGTGCTGAACCTGTGTGCAAGGAATGGTAC 540
 Qy 181 AtgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 541 CGASTGACCTCTGATGGCATGTGTGGAAAGAGCTTATCGAGAGAAATGTCAGGACAGAT 600
 Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrrLeuPheLysAsnLys 220
 Db 601 TCTCTGTGAGAGGCTGGCAGACGAAGAGGATGGGACACAGTATTTATTCAAAAACAA 660
 Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrrArgAlaLeuTyrrProLysIleIle 240
 Db 661 CTTCTGTGAGCGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCTTAAATATATA 720
 Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 Db 721 CRAGACATTGAGACATAGAAATCTAATTGGAGATGCTGGAAGACATAGTTTACAGAGAATT 780
 Qy 261 HisCysArgSerGluThrSerLysGlyValTyrrCysLeuGlnTyrrAspAspGlnLysIle 280
 Db 781 CACTGCCGAAGTCAACCAAGCAAGAGAGGTTTACTGTTTACAGATATGATGATCAGAAAAATA 840
 Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleLeuTrpAspLysAsnThrLeuGluCysLys 300
 Db 841 GTAAGCGGCTTCGAGACCAACAATCAAGATCTGGGATAAAAACACATTTGGAATGCAAG 900
 Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrrAspGluArgValIle 320
 Db 901 CGAATTTCTCAGAGCCATACAGGTTCACTCTCTCTCCAGTATGATGAGAGAGTGATC 960
 Qy 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 Db 961 ATAACAGAGATCATCGGATTCACGGTCAGAGTGGGATGTAAATACAGGTGAAATGTA 1020
 Qy 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 Db 1021 AACACGTTGATTCACCATTTGTGAAGCAGTCTCGCACTTCGCTTCAATGTTGTAGACTTTGATGAC 1080
 Qy 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 Db 1081 GTACCTGTCTCCAAAGATCGTTCCATTCGTGTATGGGATATGGCTCCCCAACCTGACATT 1140
 Qy 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 Db 1141 ACCCTCCGAGGGTGTGCTGTCGACACCCGAGCTGCTGTCAATGTTGTAGACTTTGATGAC 1200
 Qy 401 LysTyrrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 Db 1201 AAGTACATGTTTCTGCACTCGGGATAGAACTATTAAGGTATGGAACACAAAGTACTTGT 1260
 Qy 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrrArgAsp 440
 Db 1261 GAATTTGTAAGGACCTTAAATGGACACAAACCGAGGCAATTCCTGTTTGCAGTACAGGGAC 1320
 Qy 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 Db 1321 AGSCTGGTAGTGTAGTGGCTCATCTGCAACACCTATCAGATTATGGACATAGAAATGTGT 1380
 Qy 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
 Db 1381 GCATGTTTACAGCTGTAGAGGCCATGAGGAATTTGGTCGTTGATTTCGATTTGATTAAC 1440
 Qy 481 LysArgIleValSerGlyAlaTyrrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
 Db 1441 AAGAGGATAGTCACTGGGCGCTATGATGAAAAATTAAGTGTGGATCTTGTGGCTGCT 1500
 Qy 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 Db 1501 TTGGACCCCGTCTCTCTCGAGGAGCACTCTGTCTACGAGCCCTTGTGGAGCATTCGGA 1560

521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspThr 540
 1561 AGAGTTTTTCGACTACAGTTTGAGAAATCCAGATTGTCCAGTAGTTCCATGATGACACA 1620
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1621 ATCTCATCTGGGACTTCCTAAATGATCCAGCTGCCAAGCTGAACCCCGGTTCCCT 1680
 561 SerArgThrThrThrThrIleSerArg 569
 1681 TCTCGAACATACACCTACATCTCCAGA 1707

2

1 X86501 standard; cDNA; 2151 BP.

X86501;

-MAR-2003 (updated)

-SEP-1999 (first entry)

NA encoding a human beta-transducin repeat containing protein.

ta-transducin repeat containing protein; beta-Trcp; Skplp;
 osteosome degradation pathway; Vpu protein; beta-catenin;
 man immune deficiency virus-1; HIV-1; cellular protein; IkappaB;
 ubiquitinylation; phosphorylated protein; tumour; apoptosis; Alzheimer's;
 viral; antitumour; cell cycle regulation; protein degradation;
 d anti-inflammatory; osteo-articular inflammation; acute inflammation;
 mour necrosis factor; ss.

to sapiens.

Y Location/Qualifiers

70..1779

/*tag= a

/product= beta-Trcp

/note= "beta-transducin repeat containing protein"

938969-A1.

-AUG-1999.

JAN-1999; 99WO-FR00196.

-DEC-1998; 98FR-0015545.

-JAN-1998; 98FR-0001100.

ISP) INST PASTEUR.

IRM) INST NAT SANTE & RECH MEDICALE.

inzana Seisdedos F, Benarous R, Concordet J, Durand H;

all M, Margottin F;

; 1999-469329/39.

SDB; AAY24054.

, human beta-transducin repeat containing protein and its
 gments useful as, or to screen for, antiviral, antitumour,
 i-inflammatory and anti-Alzheimer's agents

im 7; Page 57-60; 71pp; French.

present sequence encodes a human beta-transducin repeat containing
 tein (beta-Trcp). The protein directs proteins to the proteosome
 radiation pathways. The protein is able to interact with the Vpu
 tein of human immune deficiency virus-1 (HIV-1), cellular proteins
 ppaB or beta-catenin (bc) and/or protein Skplp. The protein controls
 quitylation of phosphorylated proteins and thus their targeting to
 teosomes for degradation. Depending on whether the process is
 inhibited or promoted, the result may be delayed breakdown of CD4 (in
 es of HIV-1 infection); increased activity of IkB (and thus reduced
 ivity of NFkappaB) and increased degradation of mutant bc in tumour

CC cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
 CC patients. The beta-Trcp protein, and its active peptide fragments, or its
 CC nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
 CC antitumour agents that disrupt cell cycle regulation or protein
 CC degradation in human tumour cells, and anti-inflammatory agents that
 CC disrupt activation by NFkappaB. Fragments of the protein are also
 CC useful for treating osteo-articular inflammation or acute inflammation
 CC associated with release of tumour necrosis factor.
 CC (Updated on 20-MAR-2003 to correct PA field.)
 XX

SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Alignment Scores:

Pred. No.: 5,66e-303 Length: 2151
 Score: 3034.00 Matches: 569
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0

US-09-601-168B-2 (1-569) x AAX86501 (1-2151)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 70 ATGGACCGCGCGGCGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
 Qy 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
 Db 130 AGAAGAAGCTGTAATAATGCGAACCCCTAGGAAGATAATAACAGAGAGAAATTCACCT 189
 Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAACACAGTATGTTAGCA 249
 Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 250 AGCACTGCTATGAAGACTGAGAATTGTGTGGCCMAAACAAACTTGCCAATGGCACTTC 309
 Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 310 AGTATGATTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGAGAAAGGAAGT 369
 Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 370 TGTGTCAATACTTTGAGCAGTGTGCAGAGTCAGATCAAGTGGAAATTTGTGGAACATCT 429
 Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 430 ATATCCCAATGTGTCAATACCAACATGGGCAATAAACTCGTATCTTAAACCTATGTTG 489
 Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 490 CAGAGAGATTTCATACTGCTCTGCCAGCTCGGGGATGGATCATATCGTGAGAACATT 549
 Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 550 CTGTCTACCTGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTGCAAGGAAATGTTAC 609
 Qy 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAGAGCTTATCGAGAGAAATGGTCAGGACAG 669
 Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTTPGlyGlnTyrLeuPheLysAsnLys 220
 Db 670 TCTCTGTGAGAGGCTGCGACAGACGAGAGATGGGACAGTATTTATTCAAAAACAAA 729
 Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 Db 730 CCTCTCTGACGGGAATGCTCCTCCCAACTCTTTTATAGAGCACTTTATCCTAAATATA 789
 Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 Db 790 CAAGACATTGAGACAATAGATTCTAATTGGAGATGTGGAAGACATAGTTTACAGAGAATT 849

261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspSerGlnLysIle 280
 |||||
 850 CACTGCCGAAGTGAACCAAGCAAGAGGAGTTACTGTTTACAGTATGATGATCAGAAATA 909
 |||||
 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
 |||||
 910 GTAAGCGGCTTCGAGACACACACATCAAGATCGGATTAACACACATTCGATGCAG 969
 |||||
 301 ArgIleuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 |||||
 970 CGAATTCTCAGAGCCATACAGGTTTCAGTCTCTGCTCCAGTATGATGAGAGATGATC 1029
 |||||
 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
 |||||
 1030 ATACAGGATCATCGGATTCACGGTCAGAGTGGGATGTAATACAGGTGAATGCTA 1089
 |||||
 341 AsnThrIleuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 |||||
 1090 AACACGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTTGGCTTCAATAATGGCATGATG 1149
 |||||
 361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
 |||||
 1150 GTGACCTGCTCCAAAGATCGTTCATTCCTGTTGTTGGATATGGCTCCCACTGACAT 1209
 |||||
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 |||||
 1210 ACCCTCGGAGGTGCTGTTGCGACACCGAGCTGCTCAATGTTGTAGACTTTGTATGAC 1269
 |||||
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420
 |||||
 1270 AAGTACATGTTCTGATCTGCGGATAGAACATTAAGGTATGGAAACACAACTACTTGT 1329
 |||||
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 |||||
 1330 GAATTTGAGGACCTTAATGGACACAACAGGACATTCCTGTTTGCAGTACAGGAC 1389
 |||||
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460
 |||||
 1390 AGGTGTGTAGTGTGCTCATCTGACAACTATCAGATTATGGGACATAGATGTGT 1449
 |||||
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 |||||
 1450 GCATGTTTACAGTGTAGAGGCCATGAGGAATGTTGCTGTTGTTTCAATTGATAC 1509
 |||||
 481 LysArgIleValSerGlyAlaTyrAspGlyIleLysValTyrAspLeuValAlaAla 500
 |||||
 1510 AAGAGGATAGTCACTGGGGCTATGATGGAAAAATTAAAGTGTGGGATCTTGTGGCTGCT 1569
 |||||
 501 LeuAspProAspAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 |||||
 1570 TTGACCCCTGCTCTCGAGGGACACTCTGTCTACGGACCTTGTGGAGCATTCGGA 1629
 |||||
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 |||||
 1630 AGAGTTTTTCAGTACATGTTGATGAAATTCAGATTGTCAGTATTCATATGATGACACA 1689
 |||||
 541 IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 |||||
 1690 ATCTCATCTGGGACTTCTTAAATGATCCAGCTGCCCAAGCTGAACCCCCCTTCCCT 1749
 |||||
 561 SerArgThrTyrThrTyrIleSerArg 569
 |||||
 1750 TCTCGACATACACTACATCTCCAGA 1776
 }
 }
 451229 standard; DNA; 2151 BP.
 451229;
 SEP-2000 (first entry)

Human beta-TrCP coding sequence.

E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor; nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral; anti-inflammatory; immunosuppressive; cytostatic; ds.

Homo sapiens.

Key Location/Qualifiers
 CDS 70..1779
 /*tag= a
 /product= Human_TrCP

WO200034447-A2.

15-JUN-2000.

10-DEC-1999; 99WO-US29371.

10-DEC-1998; 98US-021060;
 (SIGN-) SIGNAL PHARM INC.
 (YISS) YISSUM RES & DEV CO.

Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A; Lavon I, Yaron A;
 WPI; 2000-431294/37.
 P-PSDB; AAY96697.

Polypeptide enhancing phosphorylated IkapabB ubiquitination useful for treating disorder associated with NF-kappaB activation e.g. cancer, comprising amino acid sequence of human E3 ubiquitin ligase or its variant

Disclosure; Page 72; 77pp; English.

This DNA encodes human beta-TrCP, an F-box/WD protein family member, which has been shown to have homology to human E3 ubiquitin ligase (E3). E3 enhances ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the ubiquitin pathway is useful for identifying modulators of this process for use in treating diseases associated with activation of NF-kappa-B. In vitro analysis suggests that deletion of the F-box results in a protein that functions as a dominant negative molecule in vivo. Transient over-expression of delta-beta-TrCP (a deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells, resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat inflammatory diseases, autoimmune diseases, cancer and viral infections.

Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other:

Alignment Scores:
 Pred. No.: 5,66e-303 Length: 2151
 Score: 3034.00 Matches: 569
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-601-168B-2 (1-569) x AAA51229 (1-2151)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 |||||
 Db 70 ATGACCCCGCCGAGCGGTGCTGCAAGAGAGAGGCACCTCAAGTTTATGAATTCCTCAGAG 129
 |||||
 Qy 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
 |||||
 Db 130 AGAAGACATGTATTAATGGCGAACCCCTTAGGAGATTAATACAGAGAGATTCACCT 189
 |||||
 Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 |||||

```

190 AGACACATACAAACAGCTGTGCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCA 249
61 SerThrAlaMetLysThrGluAsnCysValAlaIleLysThrLysLeuAlaAsnGlyThrSer 80
250 AGACATGCTATGAAGACTGAGAAATTGTGTGGCCAAACAAACCTTGCCCAATGGCACTTCC 309
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysLeu 100
310 AGTATGATTGTGCCCAAGCAAGCAAGCTCTCAGCAAGCTATGAAAGGAAAGCAAGCTG 369
101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
370 TGTGTCAATACTTTGAGCAGTGTGTCAGAGTCAGATCAAGTGAATTTGTGGAACATCTT 429
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
430 ATATCCAAATGTGTCATTACCAACATGGGCACATAAATCTCGTATCTTAAACCTATGTTG 489
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
490 CAGAGAGATTTCATACTGCTGCGCAGCTCGGGGATTGGATCATATCGCTGAGAACATT 549
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaIleGluLeuValCysLysGlnTrpTyr 180
550 CTGTCATACCTGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTGCAAGAAATGGTAC 609
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
610 CGAGTGACCTCTCATGGCATGCTGTGGAAGAGCTTATCGAGAGAAATGTCACGACAGAT 669
201 SerLeuTrpArgGlyLeuAlaGluArgAlaGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
670 TCTCTGTGGAGAGCCTGCGACAAACGAAGAGNTGGGACAGTATTATTCAAACCAAA 729
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
730 CCTCCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACTTATCTCTAAAAATTATA 789
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
790 CAAGACATTGAGCAATAGAAATCTAATTGGAGATGGAGACATAGTTTACAGAGAATT 849
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
850 CACTGCGGAAGTGAACAAAGCAAGAGTTTACTGTTACAGTATGATGATCAGAAATA 909
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
910 GTAAGCGGCCCTTCGAGACAAACAATCAAGATCTGGGATAAAACACATTTGGAATGCAAG 969
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
970 CGAATTCACAGGCCATACAGGTCAGTCTCTGCTCCAGTATGATGAGAGATGATC 1029
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
1030 ATAACAGGATCATCGGATTCACCGGTCCAGAGTGTGGGATGTAATACAGGTAATGCTA 1089
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1090 AACAGGTTGATTCAACATTGTGAAGCAGATTCTGCATCTGGGTTTCATTAATGGCATGATG 1149
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1150 GTGACCTGCTCCAAAGATCGTTCATTCGTGTGTATGGGATATGGCTCCCACTGACATT 1209
381 ThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1210 ACCCTCCGAGGGTGTGGTCGACACCGAGCTGCTGTCAATGTTGTAGACTTTGATGAC 1269
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420

```

```

1270 AAGTACATTGTTTCTGCATCTGGGATAGAACTATAAAGGTATGAAACACAAAGTACTTGT 1329
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1330 GAATTTGTAAAGGACCTTAATATGGACACAAACGAGGCATTGCTGTTTGCAGTACAGGGAC 1389
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1390 AGCGTGGTAGTGAGTGGCTCATCTGACAACTATCAGATTATGGGACATAGAAATGTGT 1449
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1450 GCATGTTTACAGAGTGTAGAGGCCCATGAGGAATTTGGTGGCTGTTATTCGATTTGATAAC 1509
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
1510 AAGAGGATAGTCAGTGGGCGCTTATGATGAAAAAATTAAGTGTGGGATCTTTGTGGCTGCT 1569
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1570 TTGGACCCCGTCTCTGACGGGACACTCTGTCTACGGACCTTGTGGAGCATTCGGA 1629
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspThr 540
1630 AGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTTCAGTAGTTCATCATGATGACACA 1689
541 IleLeuIleTrpAspPheLeuAsnAspProAlaIleGlnAlaGluProProArgSerPro 560
1690 ATCCTCATCTGGGACTTCTTAATGATTCAGCTGCCAAGCTGAACCCCCCGTTCCTCT 1749
561 SerArgThrTyrThrTyrIleSerArg 569
1750 TCTCGAACATACACCTACATCTCCAGA 1776

RESULT 4
AAZ93350
ID AAZ93350 standard; cdNA; 2151 BP.
XX
AC AAZ93350;
XX
XX 16-AUG-2000 (first entry)
XX
XX Sequence encoding F-box protein FBP-1.
XX
XX F-box protein; FBP; diagnosis; treatment; screening; agonist;
XX antagonist; proliferative disorder; differentiative disorder;
XX breast cancer; prostate cancer; ovarian cancer; cancer;
XX small cell lung carcinoma; immune disorder; cardiovascular disorder;
XX inflammatory disorder; human; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 70..1779
XX /tag= a
XX /product= FBP-1
XX
XX WO200012679-A1.
XX
XX 09-MAR-2000.
XX
XX 27-AUG-1999; 99WO-US19560.
XX
XX 28-AUG-1998; 98US-0098355.
XX 03-FEB-1999; 99US-0118568.
XX 15-MAR-1999; 99US-0124449.
XX
XX (UUNY ) UNIV NEW YORK STATE.
XX
XX Chiaux DS, Pagano M, Latres E;
XX
XX WPI; 2000-256635/22.
XX
XX P-PSDB; AAY83041.

```

ovel nucleic acid for screening compounds useful for treating proliferative and differentiative disorders such as cancer and immune disorders comprises sequences encoding ubiquitin ligases - disclosure; Figure 3b, 3c, 3d; 245pp; English.

ucleic acids encoding substrate-targeting subunits of ubiquitin ligases with F-box motifs (F-box proteins) are useful for diagnosis of proliferative and differentiated related disorders by measuring 3P gene expression. Cells expressing such proteins or their fragments are useful for screening compounds. The compounds are agonists or antagonists, which are useful for treating a proliferative or differentiative disorder in a mammal such as breast, ovarian and prostate cancer and small cell lung carcinoma and also major opportunistic infections, immune disorders, cardiovascular diseases and inflammatory disorders. FBP protein, analogs, derivatives and their subsequences, anti-FBP antibodies are also useful in diagnosis of the disorders.

quence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

nt Scores:

O.: 5.66e-303 Length: 2151
 3034.00 Matches: 569
 Similarity: 100.00% Conservative: 0
 cal Similarity: 100.00% Mismatches: 0
 atch: 100.00% Indels: 0
 21 Gaps: 0

01-168B-2 (1-569) x AAZ93350 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 70 ATGACCGCGCGAGGGCGGTGCTGCAAGAGAGAGCGCACTCAAGTTTATGATTCCTCAGAG 129
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
 130 AGAGAAGACTGTATAATATGCGAAGCCCTAGGAAGATAATACAGAGAGAAATTCACATT 189
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAACCAAGAACAGATGATTTAGCA 249
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 250 AGCACTGCTATGAAGACTGAGAAATGTGTGCCAAAAACAAACTTGCCAAATGSCACTTCC 309
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 310 AGTATGATTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAGAAAGAACTG 369
 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 370 TGTGTCAAAATCTTTAGCAGTGTCTCAGAGTCAGATCAAGTGGAAATTTGTGGAACATCT 429
 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 430 ATATCCCAATGTGTCTATACCAACATGGGCACATAAATCGTATCTTAAACCTATGTGTG 489
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 490 CAGAGAGATTTTCATAACTGCTCTGCCAGCTCGGGATTTGGATCATATCGCTGAGAAT 549
 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 550 CTGTCATACCTGGATGCCAAATCACTATGTCTCTCTGCTCTGCTGCTGCAAGGAATGGTAC 609
 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 610 CGATGAGCTCTGATGGCATGCTGTGGAAGAGCTTATCAGAGAAATGGTCAGGACAGAT 669
 201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220

670 TCTCTGTGGAGAGCTTGGCAGAACGAGAGGATGGGACAGATATTTATTTCAAAACAAA 729
 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 730 CCTCTGTACGGGAATGCTCTCTCCAACTCTTTTATAGAGCACTTTATCTTAAATTTATA 789
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 790 CAAGACATTCAGACAATAGAAATCTAAATGGAGATGGGAAGACATAGTTTACAGAGAATT 849
 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 850 CACTGCCAAGTGAACAGCAAGAGAGGTTTACTGTTTACATGATGATGATCAGAAATA 909
 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 910 GTAAGCGGCTTCCAGACAACACAAATCAAGATCTGGGATAAAACACACATTGGAAATGCAAG 969
 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 970 CGAATTCACAGGCCCATACAGGTTCACTCTCTCTCCAGTATGATGAGAGAGTGATC 1029
 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 1030 ATAACAGGATCATCGGATTCACCGGTCCAGATGTGGATGTAAATACAGGTGAAATGCTA 1089
 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 1090 AACACGTTGATTCAACATTTGGAAGCAGTTCTGCACCTTCGCTTTCAATTAATGGCATGATG 1149
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1150 GTGACCTGCTCCAAAGATCGTTCCATTTGTGTATGGGATATGSCCTCCCAACTGACATT 1209
 381 ThrLeuArgGluValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 1210 ACCCTCGGAGGTGTGTGCGACACCGAGCTGCTGTCATGTTGTGTAGACTTTGATGAC 1269
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1270 AAGTACATTTCTTCTGCACTCTGGGATAGACATATAAAGGTATGGACACAAAGTACTTCT 1329
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 1330 GAATTTCTAAGGACCTTAAATGGACACAAACGAGGCAATGSCCTGTTTGCAGTACAGGGAC 1389
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1390 AGGCTGTGTAGTGTGCTCATCTGCACACACTATCAGATTATGGACATAGATGTGT 1449
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACAGTGTGTAGAAAGCCCATGAGGAATTTGGTGTGTTGATTCGATTGATAAC 1509
 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
 1510 AAGAGGATATGCTAGTGGGCTATGATGAGAAAATTTAAAGTGTGGATCTTGTGTGCTGT 1569
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGACCCCGTGTCTCTCGCAGGACACCTGTGTCACGACCCCTGTGGAGCATTCGCGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTTTCGACTACAGTTTGTATGAATTCAGATTGTCCAGTAGTTTCATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1690 ATCTCTATCTGGGACTTCTTAATGATCCAGCTGCCAGCTGAACCCCCCTTCCCT 1749
 561 SerArgThrTyrThrTyrIleSerArg 569

1750 TCTCGACATACACCTACATCTCCAGA 1776

5

0

1Z93710 standard; DNA; 2151 BP.

1Z93710;

-AUG-2000 (first entry)

-box protein hBetaTrCp coding sequence.

iquitin ligase; SCF; F-box protein; targeted degradation;
 stabilisation; proteolysis; drug discovery; gene therapy; cancer;
 icoproten; Huntington's disease; gene knockout; delivery systems;
 man; ss.

no sapiens.

Y Location/Qualifiers

S 70..1779

/*tag= a

/product= F-box protein hBetaTrCp

200022110-A2.

-APR-2000.

-OCT-1999; 99WO-US23705.

*OCT-1998; 98US-0103787.

ARD; HARVARD COLLEGE.

ou P, Howley P;

I; 2000-317970/27.

PSDB; AAY83250.

rgeting degradation of polypeptide useful for treating cancer and
 her proliferative disorders, involves conjugating polypeptide with
 iquitin protein ligase or inhibiting ubiquitination using organic
 upound

aim 10; Page 171; 185pp; English.

e F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 gasses) which can be used for the targetted degradation of a target
 lypeptide in vivo. Targetted degradation is achieved by expressing
 e ubiquitin ligase in a cell linked to the interaction domain of
 e target polypeptide and thereby recruiting the target polypeptide
 the ubiquitin ligase. Such methods are useful for decreasing or
 reasing the level of a target polypeptide and for creating and
 reasing a destabilized polypeptide which is subjected to SCF
 iated proteolysis. Degrading any desired protein in a cell is
 fful for preventing or treating diseases caused by the presence of
 normal amount of the specific polypeptides, for drug discovery and
 r gene therapy. Diseases treated include cancer, by degradation of
 roproteins, Huntington's disease, other proliferative disorders and
 robial infections. The method provides a quick and easy
 ernative to gene knockout technology. The target polypeptide can
 degraded at all stages, or a specific stage, of development in the
 .ure animal.

quence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

it Scores:

Length:	2151
Matches:	569
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	21

```

US-09-601-168B-2 (1-569) x AAZ93710 (1-2151)
QY      1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db      |||||
70 ATGACCCCGCCGAGCGGCTGCTGCAAGAGAGGCACTCAAGTTATGAAATTCCTCAGAG 129
QY      21 ArgGluAspCysAsnAsnGlyGluProProArgLysLysLysLysLysLysSerLeu 40
Db      |||||
130 AGAGAGAGACTGTAATAATGCGGAACCCCTAGAGAGATAATACCAGAGAGAAATTCACCT 189
QY      41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlnThrValCysLeuAla 60
Db      |||||
190 AGACAGACATACACAGCTGTGCCAGACCTCTGCTTAAACCAAGAACAGTATGTTAGCA 249
QY      61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db      |||||
250 AGCACTGCTATGAAGACTGAGAATTGTGTGGCCAAACAAAACTTGCCAATGGCACTTCC 309
QY      81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db      |||||
310 AGTATGATTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAAGGAAAGAACTG 369
QY      101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db      |||||
370 TGTGTCAAAATACCTTTGAGCAGTGGTCAGAGTCAGATCAAGTGAATTTGTGGAACATCTT 429
QY      121 IleSerGlnMetCysHisTyrGlnHisGlyHisLysLysLysLysLysLysProMetLeu 140
Db      |||||
430 ATATCCAAATGTGTCATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 489
QY      141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisLysLysLysLysLys 160
Db      |||||
490 CAGAGAGATTTCATTAACCTGCTCTGCCAGCTCGGGAGTTGATCATATCGCTGAGAACAT 549
QY      161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
Db      |||||
550 CTGTCTATACCTGGATGCCAAATCACTATGTGCTGTCTGAACTGTGTGCAAGGAATGGTAC 609
QY      181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuLysGluArgMetValArgThrAsp 200
Db      |||||
610 CGAGTGACCTCTGATGGCATGTGTGGAGAGAGCTTATCGAGAGAAATGGTCAGGACAGAT 669
QY      201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
Db      |||||
670 TCTCTGTGGAGAGGCTTGGCAGAACCGAGAGGATGGGACAGTATTTATTTCAAAACAAA 729
QY      221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysLysLys 240
Db      |||||
730 CCTCCTGACGGGAATGCTCTCTCCAACTCTTTTATAGAGCACTTTATCTCTAAATATA 789
QY      241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db      |||||
790 CAAGACATTCAGACATAGATATCTAATTGGAGATGTGGAGACATAGTTTACAGAGAAAT 849
QY      261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysLys 280
Db      |||||
850 CACTGCCGAAGTGAACCAAGCAAGAGGAGTTACTGTGTTTACAGTATGATGATCAGAAAAA 909
QY      281 ValSerGlyLeuArgAspAsnThrIleLysLysLysLysLysLysLysLysLysLysLys 300
Db      |||||
910 GTAAGCGGCTTCGAGACAAACCAATCAAGATCTGGGATGAAACACACATTTGGAATGCAAG 969
QY      301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGlnArgValIle 320
Db      |||||
970 CGAATTCACAGGCCCATACAGGTTCCAGTCTCTGCTCCAGTATGATGAGAGAGTGATC 1029
QY      321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db      |||||
1030 ATAACAGAGATCATCGGATTCACCGGTGAGAGTGTGGGATGTAATATACAGGTGAAATGCTA 1089
QY      341 AsnThrLeuIleHisLysCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db      |||||

```


1090 AACAGGTGATTCACCATTTGTGAAGCAGTTCTGCAGCTTGGTTTCAATATGCGCATGATG 1149
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1150 GTGACCTGCTCCAAAGATCGTTCATTGCTGTATGGGATATGGCCCTCCCAACTGACATT 1209
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 1210 ACCCTCGGAGGTGCTGTGCGACACCGAGCTGCTGCTCAATGTTGTAGACTTGTATGAC 1269
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1270 AAGTACATTTGTTCTGCATCTGGGATAGAACTATAAAGGTATGGAAACAACAAGTACTTGT 1329
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 1330 GAATTTGTAGAGACCTTAATATGGACAAACAGAGGCAATCCCTGTTTGAGTACAGGGAC 1389
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1390 AGGCTGTAGTGTGCTCATCTGCACACACTATCAGATTATGGGACATAGATGTGT 1449
 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACAGATGTTAGAGGCCATGAGGAATGGTGGCTGTATTTCGATTGTATAC 1509
 481 LysArgIleValSerGlyValAlaTyrAspGlyLysIleLysValTrpAspLeuValAla 500
 1510 AAGAGGATAGTCACTGGGGCTATGATGGAAATTAAGTGTGGATCTTGTGGCTGTCT 1569
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGGACCCCGTGTCTCTGAGGAGACACTCTGTCTACGGACCCCTTGTGGAGCATTCGGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTCGACTACAGTTTGATGAATTCAGATTGTCAGTAGTTCATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1690 ATCTCTATCTGGGACTCTCTAAATGATCCAGCTGCCCAAGCTGAACCCCGGCTCCCT 1749
 561 SerArgThrTyrThrIleSerArg 569
 1750 TCTCGAATACACCTACATCTCCAGA 1776

6
0

C84610 standard; DNA; 2151 BP.

C84610;

-APR-2001 (first entry)

man ZF11 protein encoding DNA.

phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
 3; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
 Bcl-2; tumour; cytosolic; ds.

no sapiens.

200075184-A1.

-DEC-2000.

-JUN-2000; 2000WO-US15449.

-JUN-1999; 99US-0137494.

fYA) UNIV YALE.

ang H, Tsvetkov LM, Kondo T;

XX WPI; 2001-061703/07.
 DR P-PSDB; AAB48298.
 XX
 PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
 involves altering levels of proteins such as S-phase kinase associated
 proteins 1, 2 and cullin/CDC53 proteins -
 XX
 PS Examples; Page 129-130; 162pp; English.
 XX
 CC The invention relates to methods of altering the polypeptide levels in a
 cell, using proteins selected from S-phase kinase associated proteins 1
 and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
 cullin/ CDC53 family of proteins). The method is useful for altering the
 level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
 polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
 detecting tumours, and in monitoring tumor treatment in a mammal. Agents
 that modulate interactions between SKP and target proteins are useful for
 treating tumours.
 CC
 XX Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Alignment Scores:

Pred. No.: 5,66e-303 Length: 2151
 Score: 3034.00 Matches: 569
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-601-168B-2 (1-569) x AAC84610 (1-2151)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 70 ATGACCCCGCCGCGCGGTGCTGCAAGAGAGGCACTCAAGTTATGATTCCTCAGAG 129
 Qy 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
 Db 130 AGAAGAACACTGTATTAATGCGCAACCCCTAGAGAGTAATACCAGAGAGAAATTCACCT 189
 Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 190 AGACAGACATACACACAGCTGTGCCAGACTCTGCTTAAACCAAGAACACATATGTTAGCA 249
 Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 250 AGCACTGCTATGAAGACTGAGAAATTGTGTGGCCAAACAAACTTGCCCAATGCGCACTTC 309
 Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 310 AGTATGATTGTGCCAAGCAACGGAACCTCTAGCAAGCTATGAAAGGAAAGGAACTG 369
 Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 370 TGTGTCAAATACITTTGAGCAGTGTGTGAGATCAAGTGAATTTGTGGAACATCTT 429
 Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 430 ATATCCAAATGTGTCATTACCAACATGGGCACTAAACTCGTATCTTTAAACCTATGTTG 489
 Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 490 CAGAGAGATTTTCATNACTGCTCTGCCAGCTCGGGGATTGATCATATCGCTGAGAACAATT 549
 Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 550 CTGTCTACCTGATGCCAAATCATTATGTGCTGCTGAACCTGTGTGCAAGGAATGGTAC 609
 Qy 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 610 CGAGTGACCTCTGATGGCATGCTGTGGAGAGAGCTTTATCGAGAGAAATGGTCAGACAGAT 569
 Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220

```
670 TCTCTGGAGAGCGCTGGCAGAACGAAGAGGATGGGACAGTATTTATTCAAAACAAA 729
|||||
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
|||||
730 CCTCTGACGGGAATGCTCTCCCACTCTTTTATAGAGCACTTATCTCAAAATTATA 789
|||||
241 GluAspIleGluThrIleGluSerAsnThrArgCysGlyArgHisSerLeuGlnArgIle 260
|||||
790 CAGACATTGAGACAATAGAAATCTAATTCGAGATGTGGAAGACATAGTTTACAGAGAATT 849
|||||
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
|||||
850 CACTGCCGAAGTGAACAACGAAGAGAGTTACTGTTTACAGTATGATCATCAGAAAATA 909
|||||
281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
|||||
910 GTAAGGGCCTTCGAGACACACAATCAAGATCTGGGATAAAAAACACATTGGAATGCAAG 969
|||||
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
|||||
970 CGAATTCTCAGCGCCATACAGGTTGAGTCTCTGCTCCAGTATGATGAGAGAGTGATC 1029
|||||
321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
|||||
1030 ATACAGGATCATCGGATTCACGGTCCAGGTGTGGATGTAAATACAGGTGAATGCTA 1089
|||||
341 AsnThrIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
|||||
1090 AACACGTTGATTACCACTTTGTGAAGCAGTTCTGCACTTGGTTTCAATTAATGGCATGATG 1149
|||||
361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
|||||
1150 GTACCTGCTCCAAAGATCTTCCATTGCTGTGATGGATATGGCTTCCCACTGACATT 1209
|||||
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
|||||
1210 ACCCTCGGAGGGTGTGTCGACACCGAGTGTCTCAATGTTGTAGACTTTGATGAC 1269
|||||
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420
|||||
1270 AAGTACATTGTTTCTGCATCTGGGGATAGAACTATAAAGGTATGGAACCAAGTACTTGT 1329
|||||
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
|||||
1330 GAATTTGTAGGACCTTAATAGCACACAACAGGGCATTCCTGTTTGCAGTACAGGAC 1389
|||||
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460
|||||
1390 AGGCTGGTGTAGTGGCTCATCTGACACACACTATCAGATTATGGGACATAGAAATGTG 1449
|||||
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
|||||
1450 GCATGTTTACAGTGTAGAAAGCCATAGAGAAATGGTGGTGTGTTATTCGATTGAAAC 1509
|||||
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTyrAspLeuValAlaAla 500
|||||
1510 AAGAGGATAGTCACTGGGGCTATGATGGAATAATTAAGTGTGGATCTTGTGGCTGCT 1569
|||||
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
|||||
1570 TTGACCCCGGTGCTCTCGAGGGACACTCTGTCTACGGACCCCTTGTGAGCATTTCCGGA 1629
|||||
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
|||||
1630 AGAGTTTTCAGCTACAGTTTGTAGTTCAGATTTCAGTASITTCATGATGCACACA 1689
|||||
541 IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
|||||
1690 ATCTCATCTGGGACTTCTTAATGATCCAGCTGCCAAGCTGAACCCCGGCTTCCCT 1749
|||||
561 SerArgThrTyrThrTyrIleSerArg 569
|||||
```

Db 1750 TCTCGAACATACACCTACATCTCCAGA 1776

RESULT 7

AAL41041

ID AAL41041 standard; cDNA; 2151 BP.

XX AAL41041;

AC AAL41041;

XX 11-OCT-2002 (first entry)

XX cDNA of Human F-box protein FBPI SEQ ID No 1.

XX Cytostatic; immunomodulator; cardiant; antiinflammatory; antimicrobial;

KW proliferative; differentiative disorder; Skp2; F-box protein; cancer;

KW ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;

KW small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;

KW inflammatory disorder; lymphoma; major opportunistic infection;

KW certain cardiovascular disease; human; gene; ss.

XX Homo sapiens.

OS

XX WO200255665-A2.

FN

XX 18-JUL-2002.

PD

XX C7-JAN-2002; 2002WO-US00311.

PF

XX 05-JAN-2001; 2001US-260179P.

FR

XX (UJNY) UNIV NEW YORK STATE.

PA

XX Pagano M;

PI

XX WPT; 2002-599665/64.

DR

XX P-PSDB; AAO22446.

DR

Screening compounds for treating proliferative disorders, e.g. breast cancer or prostate cancer, infections or immune disorders, comprises detecting a change in the activity of Skp2 with either p27 or Cks1 -

XX Disclosure; Fig 3; 246pp; English.

XX The invention relates to screening compounds useful for the treatment of proliferative or differentiative disorders comprising detecting a change in the activity of Skp2 (F-box protein). The method is useful for screening compounds for the treatment of proliferative or differentiative disorders, particularly cancer. These compounds include small molecules, or compounds or derivatives or analogues of the new ubiquitin ligases. The compounds are useful for treating diseases such as cancer (e.g. lung carcinoma or parathyroid adenomas), major opportunistic infections, breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell lung carcinoma or parathyroid adenomas), major opportunistic infections, immune disorders, certain cardiovascular diseases or inflammatory disorders. This polynucleotide sequence represents the cDNA encoding an F-box protein (FBP) of the invention.

XX SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Alignment Scores:

Pred No.: 5,66e-303 Length: 2151

Score: 3034.00 Matches: 569

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 24 Gaps: 0

US-09-601-168B-2 (1-569) x AAL41041 (1-2151)

QY 1 MetAspProAlaGluAlaValLeuGlnLysAlaLeuLysPheMetAsnSerSerGlu 20

Db 70 ATGACCCGCGGAGCGGTCTGCAAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129

QY 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40

```

130 AGAGAGACACTGTAATAATGCGGAAACCCCTAGGAAGATATAACAGAGAAGAAATTCACCTT 189
41 ArgGlnThrTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1270 AAGTACATGTTTCTGCATCTGGGATAGAACTATAAAGGTATGAACACAACTACTTGT 1329
190 AGACAGACATACACACAGCTGTGCGAGACTCTGCTTAAACCAAGAAACAGATATGTTAGCA 249
61 SerThrAlaMetLysThrGluAsnGlyHisLysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
250 AGCACTGCTATGAAGACTGAGAATTTGTGCGCCAAACAAACTTGCCAAATGGCACTTCC 309
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGlnLysGluLeu 460
310 AGTATGATTTGTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAAGAACTG 1449
101 CysValLysTyrPheGluGlnTyrSerGluSerAspGlnValGluPheValGluHisLeu 480
370 TGTGTCAAAATCTTTGAGCAGTGGTTCAGATCAGATCAAGTGGAAATTTGTGGAACATCTT 1509
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 500
430 ATATCCCAAAATGTCATTACCAACATGGGCACATAAATCTGATCTTAAACCTATGTTG 1569
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 520
490 CAGAGATTTTCATAACTGCTCTGCCAGCTCGGGGATTTGGATCATATCCTGAGAACATT 1629
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 540
550 CTGTCACTACCTGGATGCCAAATCACTATGTGCTGCTGAACCTTGTGTGAAGAAATGGTAC 1689
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 560
610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCAGAGAATTTGTCAGGACAGAT 1749
201 SerLeuTrpArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 569
670 TCTCTGTGGAGAGCCCTGGCAGAACGAAGAGGATGGGACAGATTTATTTCAAAACAA 1776
221 ProTrpAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 1776
730 CCTCCTCAGCGGAATGCTCTCCCACTCTTTTATAGACACTTTTATCTTAAATATATA 1776
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 1776
790 CAAGACATTGAGACAAATAGAACTAATTTGAGATGTGGAAGACATAGTTTACAGAAAT 1776
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 1776
850 CACTGCCGAAAGTGAACCAAGCAAGGAGTTTACTGTTTACAGTATGATGATCAGAAATA 1776
281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 1776
910 GTAAGCGCCTTCAGACAAACAACTCAAGATCTGGATTAACACATTTGATGATGATGATC 1776
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 1776
970 CGAATCTTCAGCGCCATCAGAGTTTCACTGCTCTGCTCAGATGATGATGAGAGATGATC 1776
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 1776
1030 ATACAGAGATCATCGGATTCACGGTCAGAGTGGGATGTAATACAGGTGAAATGCTA 1776
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 1776
1090 AACAGTTGATTCACCATTTGTAAGCAGCTTCGACCTTGGCTTCAATTAATGGCATGATG 1776
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 1776
1150 GTGACCTGCTCCAAAGATCGTTCCATTTGCTGTATGGGATATGGCTCCCACTGACATT 1776
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 1776
1210 ACCCTCCGAGGGTCTGCTGGACCGACCGAGCTGCTGCTCAATGTTGTAGATTGATGAC 1776
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1270 AAGTACATGTTTCTGCATCTGGGATAGAACTATAAAGGTATGAACACAACTACTTGT 1329
421 GluPheValArgThrLeuAsnGlyHisLysValAlaLysThrLysLeuAlaAsnGlyThrSer 440
1330 GAATTTGTAAGGACCTTAAATGGACACAAACAGGACATTCCTGTTTGCAGTACAGGGAC 1389
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1390 AGCTCGTATGTAGTGGCTCATCTGACAAACACTATCAGATTATGGGACATAGAATGCTG 1449
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1450 GCATGTTTACGAGTGTAGAGGCCATGAGGAATTTGGTGGTGTATTCGATTTCGATTAC 1509
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
1510 AAGAGGATAGTCACTGGGCTTATGATGGAATAATTAAGTGTGGATCTTGTGGCTGCT 1569
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1570 TTGACCCCGCTCTCTCGAGGACACTCTCTTACGAGCCCTTGTGAGACATTTCCGGA 1629
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
1630 AGAGTTTTCGACTACAGTTTGTGAATTCAGATTGTCAGTAGTTCAATGATGACACA 1689
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
1690 ATCTTCATCTGGGACTTCTTAATGATCCAGCTGCCAAGCTGGAACCCCGCTTCCCT 1749
561 SerArgThrTyrThrTyrIleSerArg 569
1750 TCTCGAACATACACCTTACATCTCCAGA 1776

RESULT 8
ABS51009
ID ABS51009 standard; cDNA; 2151 BP.
XX
AC ABS51009;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human cDNA encoding bait protein beta-TrCP1.
XX
KW Human; ss; gene; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
non-insulin diabetes mellitus; obesity; selected interacting domain;
XX
KW SID; protein-protein interaction map; PWM; anorectic; metabolic disorder.
XX
OS Homo sapiens.
XX
PN WO200253726-A2.
XX
PD 11-JUL-2002.
XX
PF 28-DEC-2001; 2001WO-BP15423.
XX
PR 02-JAN-2001; 2001US-259377P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI Legrain P, Marullo S, Jockers R;
XX
DR WPI; 2002-583612/62.
DR P-PSDB; ABG69473.
XX
PT Novel complex of protein-protein interactions in adipocyte cells for
identifying compounds that modulate the protein-protein interactions
PT and useful for treating obesity and metabolic disorders.
XX

```

claim 2; Page -; 125pp; English.

The invention relates to a complex of protein-protein interactions forming a protein-protein interaction map, PIM) in adipocyte cells as defined in the specification, or polynucleotides in adipocytes encoding for the polypeptides. Also included are a recombinant cell expressing the interacting polypeptides and a method of selecting a modulating compound in adipocyte cells, by culturing a recombinant host cell on a selective medium containing a modulating compound and a reporter gene the expression of which is toxic for the recombinant host cell which is transformed with two vectors, where the first vector comprises a polynucleotide encoding a first hybrid polypeptide and DNA binding domain and the second vector comprising a polynucleotide encoding a second hybrid polypeptide and an activating domain that activates the basic reporter gene, when the first and second hybrid polypeptides interact and selecting the modulating compound which inhibits the growth of the recombinant host cell (i.e. using the yeast two-hybrid system). The complexes are useful for identifying compounds that modulate protein-protein interactions and useful for treating obesity and metabolic disorders e.g. non-insulin dependent diabetes mellitus, and the compound isolated by the method is useful for treating and preventing obesity or metabolic diseases. The interactions between the proteins of the complex further define a set of selected interacting proteins, SID. The present sequence encodes a member of the protein complex of the invention, used as the bait protein in the yeast two-hybrid assay.

The present sequence was not displayed in the specification but is obtained from its Genbank entry by the indexer.

quence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

nt Scores:

O.:	5.66e-303	Length:	2151
Similarity:	100.00%	Matches:	569
cal Similarity:	100.00%	Conservative:	0
atch:	100.00%	Mismatches:	0
	24	Indels:	0
		Gaps:	0

01-168B-2 (1-569) x ABS51009 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLeuAlaLeuLysPheMetAsnSerSerGlu 20
 70 ATGGACCGCGCGAGGGGTGCTGCAAGAGAGAGCACTCAAGTTTATGAATTCCTCAGAG 129
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleLeuProGluLysAsnSerLeu 40
 130 AGAAGAAGACTGTATATGGCGAACCCTTAGGAAGATAATACAGAGAGAAATTCACCT 189
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCA 249
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLysLeuAlaAsnGlyThrSer 80
 250 AGCACTGCTATGAAGACTGAATTTGTGGCCAAACAAACTTCCCAATGGCACTTCC 309
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 310 AGTATGATTGTGCCAACGACGAAACTCTCAGCAGACTATGAAAGGAAAGGAACCTG 369
 101 CysValLysTyrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
 370 TGTGTCAAAATCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTGGAACTCTT 429
 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 430 ATATCCCAAAATGTTCATTACCAACATGGGCACATAAATCGTATCTTAAACCTATGTTG 489
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 490 CAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGATGGATTCATATCATATCGCTGAGACATT 549

QY	161	LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr	180
DB	550	CTGTCACTACCTGATGCCAATCACTATGTCTGTCTGAACTTGTGTCAAGGAATGCTAC	609
QY	181	ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValargThrAsp	200
DB	610	CGAGTGACCTCTGATGCGATGCTGTGTGAAGAAGCTTTATCGAGAGAAATGTCAGGACAGAT	669
QY	201	SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys	220
DB	670	TCTCTGTGGAGGCGCTGGCAGAACGAGAGGATGGGACAGTATTTATCAAAACAAA	729
QY	221	ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle	240
DB	730	CCTCCTCAGCGGAATGCTCTCCCAACTCTTTATATAGAGCACTTTATCTTAAATATA	789
QY	241	GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle	260
DB	790	CNAGACATTGAGACATAGATCTAATTGGAGATGTGGAGACATAGATTACAGAGATT	849
QY	261	HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle	280
DB	850	CACCTGCCGAAGTGAACCAAGAAAGGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA	909
QY	281	ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys	300
DB	910	GTAAGCGGCTTCGAGACACACAAATCAAGATCTGGGATATAAACAACATTGGAATGCAAG	969
QY	301	ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle	320
DB	970	CGAATTCACAGGCCCATACAGGTTCACTCTCTCTCTCCAGTATCAGAGAGTGATC	1029
QY	321	IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu	340
DB	1030	ATAACAGGATCATCGGATTCACCGGTGAGAGTGTGGGATGTAATACAGGTGAAATGCTA	1089
QY	341	AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet	360
DB	1090	AACAGGTTGATTCAACCATTTGGAAGCAGTTCTGCACTTTCATTAATGSCATGATG	1149
QY	361	ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle	380
DB	1150	GTGACCTGCTCCAAAGATCGTTCATTGCTGTATGGGATATGCGCTCCCACTGACATT	1209
QY	381	ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp	400
DB	1210	ACCTCCGAGGGTGTGTCGACACCGAGCTGCTGTCAATGTTGTAGACTTTGATGAC	1269
QY	401	LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys	420
DB	1270	AAGTACATTGTTCTGCACTGGGATAGAACTATAAAGGTATGGAACACCAAGTACTTGT	1329
QY	421	GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp	440
DB	1330	GAAATTTGTAAGGACCTTAAATGGACACAAACGAGGCACTTGCCTGTTTGCAGTACAGGAC	1389
QY	441	ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly	460
DB	1390	AGGCTGTGTAGTGTGCTCATCTGCAACACCTATCAGATTATGGGACATAGAACTGTGT	1449
QY	461	AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn	480
DB	1450	GCATGTTTACAGTGTGTAGAGGCCCATGAGGAATTCGGTGTGTATTCGATTGTGATAAC	1509
QY	481	LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla	500
DB	1510	AAGAGGATAGTCACTGGGGCTATGATGGAAAAATTAAGTGTGGGATCTTGTGGCTGCT	1569
QY	501	LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly	520
DB	1570	TTGACCCCGTGTCTCTCGCAGGACACTCTGTCTACGAGACCCCTTGTGGAGCATTCGGA	1629
QY	521	ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspThr	540

|||||
1630 AGAGTTTTCGACTACAGTTTGATGATTCAGATTGTCAGTAGTTTCACATGATGACACA 1689
|||||
541 lleleulletrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
|||||
1690 ATCCTCATCTGGGACTTCTCTAAATGATCCAGCTGCCCAAGCTGACCCCGGTTCCCT 1749
|||||

561 SerArgThrTyrThrTyrIleSerArg 569
|||||
1750 TCTGCAATACACTACATCTCCAGA 1776
|||||

9
33
1229233 standard; cDNA; 2419 BP.

1229233;

-FEB-2000 (first entry)

man cell signalling protein-12 encoding cDNA.

ll signalling protein-12; CSIGP-12; cell proliferation;
flammatory disorder; cirrhosis; cancer; hepatitis; AIDS;
teriosclerosis; Addison's disease; multiple sclerosis; ss.

mo sapiens.

Y Location/Qualifiers
S 70..1779
/*tag= a
/product= "Cell Signalling Protein-12"

9958558-A2.

-NOV-1999.

99NOV-1999; 99NO-US10567.

99NOV-1998; 98US-0085343.
-AUG-1998; 98US-0098010.

NCY-) INCYTE PHARM INC.

ndman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;
John MR, Yang J;

I; 2000-086432/07.
PSDB; AAY44249.

nan cell signaling proteins useful for, e.g. diagnosing cell
proliferative and inflammatory disorders

aim 9; Page 87-88; 90pp; English.

a present sequence is a cDNA obtained from Incyte clone 3239149 of
-AUCT01 library. It encodes cell signalling protein-12 (CSIGP-12). It
expressed in musculo-skeletal, gastrointestinal and nervous
issues. Fragments of CSIGP encoding nucleic acid can be used as
proliferation probe for detecting CSIGP related sequences or allelic
variants. Recombinant CSIGP can be produced in host cells by transforming
with genetically engineered vectors. Agonists or antagonists can be
used in the treatment of cell proliferative and inflammatory disorders
associated with decreased or increased CSIGP expression. CSIGP is used in
diagnosis, prevention and treatment of cell proliferative disorders
arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
disorders like AIDS, Addison's disease, multiple sclerosis, etc.

quence 2419 BP; 671 A; 531 C; 625 G; 590 T; 2 other;

it Scores:

.. 6.75e-303 Length: 2419
3034.00 Matches: 569
Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-601-168B-2 (1-569) x AA229233 (1-2419)
Qy 1 MerAspProAlaGlnAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 70 ATGGACCCGCGAGGGGCTGCGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
Qy 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
Db 130 AGAGAGACTGTAATAATGCGAACCCTCTAGAGAGATAATACAGAGAGAAATTCACCT 189
Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCA 249
Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 250 AGCACTGCTATGAGACTGAGAAATGCTGTGGCCAAACAAAACCTTGCCTATGGCACTCC 309
Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db 310 AGTATGATTGTGCCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAAAGGAAGACTG 369
Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 370 TGTGTCMAATACTTTGAGCAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTGGAACATCT 429
Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
Db 430 ATATCCCAATGTGTCTATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 489
Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGlnAsnIle 160
Db 490 CAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGATGGATCATATTGCTGAGAACATT 549
Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
Db 550 CTGTCTACCTGATGCCAAATCACTATGCTGTCTGTAACCTGTGTGCAAGGAATGGTAC 609
Qy 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
Db 610 CGAGTGCCTCTCATGSCATGCTGTGGAAGAAAGCTTATCGAGAGAAATGGTCAGACAGAT 669
Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
Db 670 TCTCTGTGGAGAGGCTTGGCAGNACGAGAGAGATGGGACAGTATTTATCAAAAACAAA 729
Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
Db 730 CCTCTGTGACGGGAATGCTCTCTCCAACTCTTTTATAGAGCACTTTATCTTAAAAATTATA 789
Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db 790 CRAACATTTCAGACATAGAACTAATTCGAGATGTCGAGACATAGTTTACAGAGAAAT 849
Qy 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
Db 850 CACTGCCGAAGTGAACCAAGCAAGGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA 909
Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
Db 910 GTTAGCGGCTTCGAGACAAACCAATCAAGATCTGGGATAAAAAACACATTGGAAATGCAAG 969
Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 970 CGAATTCACAGGCCATACAGGTTCAAGTCCCTCTCTCTCCAGTATGATGAGAGAGTGATC 1029
Qy 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 1030 ATAACAGGATCATCGGATTCACCGGTGAGAGTGGGATGTAATACAGGTGAATGCTA 1089

```

341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1090 AACACGTTGATTCACCATTTGTGAAGCAGTTCCTGCACTTCGGTTTCAATTAATGGCATGATG 1149
361 ValThrCysSerIysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1150 GTGACCTGCTCCAAAGATGCTTCATGCTGTATGGGATATGGCTCCCACTGACAT 1209
381 ThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1210 ACCCTCGGAGGTGCTGTCGACACCGAGCTGCTCAATGTTGTAGCATTTGATGAC 1269
401 LysTrpIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1270 AAGTACATTTGTTCTGCATCTGGGATAGAACTATTAAGGTATGGAACAACAAGTACTTGT 1329
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTrpArgAsp 440
1330 GAATTTGTAGGACCTTAATGACACACACGAGCATTCCTGTTTGCAGTACAGGGAC 1389
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1390 AGGCTGTGTAGTGTCTCATCTGACAACTATCAGATTATGGGACATAGAAATGCTGT 1449
461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
1450 GCATGTTACAGTGTAGAGGCCATGAGGAATGTCGTTGTATTCGATTTGATTAAC 1509
481 LysArgIleValSerGlyAlaTrpAspGlyLysIleLysValTrpAspLeuValAla 500
1510 AAGAGGATAGTCAGTGGGGCTATGATGAAAAATTAAGATGTGGGATCTTGTGGTGTCT 1569
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1570 TTGACCCCGTCTCTCGAGGACACTGTCTACGGACCTGTGTGAGGATTCGGA 1629
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
1630 AGAGTTTTCGACTACAGTTTGTATGATTCAGATTGTCAGTAGTTTCATGATGACACA 1689
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1690 ATCTCATCTGGGACTTCCTAAATGATTCAGCTGCCAAGCTGAACCCCGTTCCCT 1749
561 SerArgThrTrpThrTrpIleSerArg 569
1750 TCTCGACATACACTACATCTCCAGA 1776

10
3
490079 standard; cDNA; 3220 BP.
490079;
-OCT-2001 (first entry)
nan bone marrow cDNA, SEQ ID NO: 323.
nan bone marrow; antiinflammatory; cytostatic; neuroprotective;
viral; antibacterial; antifungal; anti-HIV; haemostatic;
nonsuppressive; gene therapy; cytokine cell proliferation;
l differentiation modulator; immune disorder; infection; cancer;
nan immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
no sapiens.
200153453-A2.
JUL-2001.
DEC-2000; 2000WO-US34960.

```

21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0662191.
19-OCT-2000; 2000US-0693036.
30-NOV-2000; 2000US-0250583.
(HYSE-) HYSEQ INC.
Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Drmanac RT;
WPI; 2001-488707/53.
P-PSDB; AAM00960.
Novel bone-marrow-expressed polynucleotides and polypeptides, useful
for treating e.g. cancer and immune deficiency disorders -
Claim 1; Page 428; 648pp; English.
The present sequence is one of 251 novel human polynucleotides
expressed in the bone marrow. The polynucleotide and the
polypeptide encoded by it are useful in the treatment of various
immune deficiencies and disorders. The deficiencies and disorders may
be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
infection, or may result from an autoimmune disorder, a coagulation
disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
suppression of an inflammatory response or treatment of a nervous
system disorder such as Alzheimer's disease. Detection of the presence
or increased expression of the polynucleotide or the protein it
encodes is useful for the diagnosis and/or prognosis of one
or more types of cancer. The polynucleotide and polypeptide can be
used as nutritional sources or supplements and in the screening of
chemical compounds as potential drugs.
Sequence 3220 BP; 878 A; 725 C; 770 G; 847 T; 0 other;
Alignment Scores:
Pred. No.: 5,46e-302 Length: 3220
Score: 3027.00 Matches: 567
Percent Similarity: 99.82% Conservative: 1
Best Local Similarity: 99.65% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 22 Gaps: 0
US-09-601-168B-2 (1-569) x AAH900079 (1-3220)
Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 128 ATGGACCGGTGCGAGGGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 187
Qy 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
Db 188 AGAGAAGACTGTAAATATGCGCAACCCCTAGCAAGATATAACCAAGAGAAGAAATTCGCTT 247
Qy 41 ArgGlnThrTrpAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 248 AGACAGCATATCAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTTAGCA 307
Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 308 AGCACTGCTATGAAGACTGAGAATTTGTGTGGCAAAACAAACTTGCCTAATGGCACTTCC 367
Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db 368 AGTATGATTGTGCCCAAGCAACGAAACTCTCAGCAAGCTATCAAAAGGAAAGGAACTG 427
Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 428 TGTCGTAATACATTTGAGCAGTGTGTCAGAGTCAGATCAAGTGGATTTGTGGAACATCTT 487

```

121 IleSerGlnMetCysHisTyrClnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
122 |||||
488 ATATCCCAAAATGTGTCAATTACCAATCGGCACATAAATCGTATCTTAAACCTATGTTG 547
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
548 CAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGATTGGATCATATTGCTGAGACATT 607
161 LeuSerTyrLeuAspAlaLysSerIenCysAlaAlaGluLeuValCysLysGluTyrTyr 180
608 CTGTCTATCTCGGATGCCAAATCAATATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
181 ArgValThrSerAspGlyMetLeuTyrLysIleLeuGluArgMetValArgThrAsp 200
668 CGAGTGACCTCTGATGGCATGCTGTGGAGAGCTTATCGAGAGATGGTCAGGACAGAT 727
201 SerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
728 TCTCTGTGGAGAGGCTGTGCAGACGAAGAGGATGGGACAGATATTATTCAAACAAA 787
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
788 CTTCTGTGAGGAGATGCTCTCCCAACTCTTTTATAGAGCACTTTATCTAAATATA 847
241 GlnAspIleGluThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIle 260
848 CAAGACATTGAGACATAGATCTAATTGGAGATGTGGAGACATAGCTTTACAGAGATT 907
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
908 CACTCCGAGGTGAACAGCAAGCAAGAGGAGTCTACTGTTTACAGTATGATGATGAGAAATA 967
281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
968 GTACGGGCTTCGAGACACACATCAAGATCTGGGNTAAACACATTTGGATGCAAG 1027
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
1028 CGAATCTCACAGGCGCATACAGGTTTCAGTCTCTGCTCCAGTATGATGAGAGATGATC 1087
321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
1088 ATAACAGGATCATCGAATTCACGCTCAGAGTGGGATGTAAATACAGGTGAATGCTA 1147
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1148 AACACGTTGATTCACCATTTGAAGCAGTTCTGCACCTGCGTTTCAATAATGGCATGATG 1207
361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
1208 GTGACCTGTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCTCCCAACTGACATT 1267
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValAspPheAspAsp 400
1268 ACCCTCCGAGGCTGCTGTCGACACCGAGCTGCTGCAATGTTGTAGACTTTGATGAC 1327
401 LysTyrIleValSerAlaSerClyAspArgThrIleLysValTyrAsnThrSerThrCys 420
1328 AAGTACATTGTTCTGGCATCTGGGATAGAACTATAAGGATGGAACACAAAGTACTTGT 1387
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1388 GAATTGTGAAGGACCTTAATAGGACACAAACGAGGCTTGCCTGTTGCGATGACAGGAC 1447
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460
1448 AGGCTGGTAGTGGTGGCTCATCTGACACACTATCAGATTATGGGACATAGATGTTGT 1507
461 AlaCysLeuArgValLeuGluClyHisGluLeuValArgCysIleArgPheAspAsn 480
1508 GATGTTTACGAGTGTGAAGGCGCATGAGGAATGGTGGCTGTTATTCGATTGATAAC 1567

```

RESULT 11

```

AAKS1715
ID AAKS1715 standard; cDNA; 2285 BP.
XX
AC AAKS1715;
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 260.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
OS Homo sapiens.
XX
FN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US040398.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
(PHYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
WPI; 2001-476283/51.
DR P-PSDB; AAM78582.
XX
Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
Claim 1; Page 1177-1180; 6221pp; English.
XX
The invention relates to polynucleotides (AAKS1456-AAKS3435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating

```

tivity, tissue growth factor activity, immunomodulatory activity and tissue inhibitor activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAK58020) are omitted as the relevant pages from the sequence listing are missing at the time of publication.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

sequence 2285 BP; 646 A; 492 C; 573 G; 574 T; 0 other;

nt scores:

4.84e-300	Length:	2285
3006.00	Matches:	569
94.05%	Conservative:	0
94.05%	Mismatches:	0
99.08%	Indels:	36
22	Gaps:	1

01-168B-2 (1-569) x AAK51715 (1-2285)

1 MetAspProAlaGluAlaValLeuGlnGlnGluValAlaLeuLys----- 14

127 ATGGACCCGGCCGAGGGCGGTGCTGCCAAGAGGCACTCAAGTTATGTGCTCTATGCC 186

14

187 AGGTCTCTGTGGCTGGGCTGCTCCAGCCTGGCGGACAGCATGCCCTTGGCTGGGATGCCCTG 246

15 -----pheMetAsnSerSerG[112]rrG[112]snGys 24

247 TATAACCCAGGCACTGGCGCACTCACAGCTTTCATGAA TTCTCAGAGAGAGAGACTGT 306

25 AsnAsnGlyGlyIleProProArgIleIleProGluIleAsnSerLeuArgGlnThrTyr 44

307 AATAATGGCGAACCCCTAGGAGGATATAATACCGAGACAGATTGGCTTACGACATTC 366

```
45  AenSexCysA] aA yC] eiuCysI eiuA enC] nC] uTh yV e] CysI eiuA] = SexTh yA] = Met 61
```

357 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

100

1

2

.....

100

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

078 069GCTCTTCAGACAGATCTCTGTGGAGA

203 GYUENATAGIAGIGYIIRGUYGINIYUEUNETYSANILYSFOPFOASBGLY 227

84 / GGGCTGGCAGAACGAGAGGATGGGACAGATTTATTCAAAACAAACCTCCTGACGGG 906

223 ASIAAFIOFIOASHISERFNEIYIARGATALEUTYIFPROLYSIELEEGINASPILEGU 244

907	Db		AATGCTCTCCCAACTCTTTTATAGAGACACTTTATCTCTAAAATTATACAGACATTGAG	966
245	Qy		ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer	264
967	Db		ACAATAGAACTCAATTGGAGATGCTGAAGACATAGTTTACAGAGAAATTCACGCGGAAGT	1026
245	Qy		GluThrSerIysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu	284
1027	Db		GAACAAGCAAGCAAGAGTTTACTGTTTACAGTATGATGATCAGAAAAATAGTAAGCGGCGTT	1086
285	Qy		ArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr	304
1087	Db		CGAGACAACAATCAAGATCTGGGATAAAAAACACATTTGGAAATCGAAGGAATTCACACA	1146
305	Qy		GlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySer	324
1147	Db		GGCCATACAGGTCAGTCCTCTCTCCCACTATGATGAGAGAGTGAATCAACAGGATCA	1206
325	Qy		SerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIle	344
1207	Db		TCSGATTCACGGTCCAGAGTGTGGGATGTAAATACAGGTGAATGCTAAACAGTTCGATT	1266
345	Qy		HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer	364
1267	Db		CACCAATTTGTGAAGCAGTCTCGACATTGCGTTTCAATAATGCGATGATGGTGACCTGCTCC	1326
365	Qy		LysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArg	384
1327	Db		AAAGATCGTTCATTCGTATGGGATATGGCTCTCCCAACTGACATTAACCTCCGGAGG	1386
385	Qy		ValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAspLysTyrIleVal	404
1387	Db		GTGCTGTGTCGGACACCGAGCTGCTGCAATGTTGTAGACTTTGATGACAAGTACATTTGTT	1446
405	Qy		SerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArg	424
1447	Db		TCTGCATCTGGGATAGACTAATAGGTATGGSACACAGTACTTTGTGAATTTGTAAAG	1506
425	Qy		ThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVal	444
1507	Db		ACCTTAAATGACACAACAGGCAATTGCCTGTTTGCAGTACAGGACAGGCTGGTAGTG	1566
445	Qy		SerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArg	464
1567	Db		AGTGCTCATCTGACACACTATCAGTTATGGGACATAGAAATGTGGTGCAATGTTTACGA	1626
465	Qy		ValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsnLysArgIleVal	484
1627	Db		GTGTTAGAAGGCCATGAGGAATTTGGTGCCTGTATTTCGATTGTATACAGAGGATAGTC	1686
485	Qy		SerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArg	504
1687	Db		AGTGGGCGCTATGATGMAAATAAAGTGTGGGATCTTTGGCTGCTTTGGACCCCGT	1746
505	Qy		AlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArg	524
1747	Db		GCTCTCGAGGCACACTCTGCTACGGACCCCTTGTGGAGCATTCGGAAGAGTTTTCGA	1806
525	Qy		LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrp	544
1807	Db		CTACAGTTTGATGAATTCACAGATTTGTCAGTAGTTTCAATGATGACACAATCCTCATCTGG	1866
545	Qy		AspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerProSerArgThrTyr	564
1867	Db		GNCTTCTTAATGATCCAGCTGCCAAGCTGAACCCCCCGCTTCCCTTCTTCGACACATC	1926
565	Qy		ThrTyrIleSerArg	569
1927	Db		ACCTACATCTCCAGA	1941

RESULT 12
AAA73131

A73131 standard; cDNA; 1707 BP.

A73131;

-NOV-2000 (first entry)

use ubiquitin ligase FWD1 protein encoding cDNA.

ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB; ta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1; ne therapy; colon cancer; beta-transducin repeat containing protein; ta-TRCP; ss.

s musculus.

2000166542-A.

-JUN-2000.

-DEC-1998; 98JP-0343437.

-DEC-1998; 98JP-0343437.

AGA-) KAGAKU GIUTSU SHINKO JIGYODAN.

I; 2000-485550/43.

PSDB; AAB12812.

box protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or beta-catenin

sclosure; Fig 16; 19pp; Japanese.

e present invention describes an F-box motif protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or ta-catenin and is constituted by Skp1 protein, Cull1 protein and a mplex (SCF complex) of F-box protein containing F-box motif and WD40 repeat motif and has the amino acid sequence of 45 residues (AAB12811) one of two 569 residue sequences (AAB12812, which is mouse ubiquitin ligase FWD1 protein) and (AAB12813, which is human beta-transducin repeat containing protein (beta-TRCP)). The F-box protein can be used for e gene therapy of colon cancer by being recombined to a virus vector. e present sequence encodes the mouse ubiquitin ligase FWD1 protein, om the present invention.

quence 1707 BP; 467 A; 399 C; 453 G; 388 T; 0 other;

nt Scores:

o.:	2,67e-299	Length:	1707
Similarity:	2997.00	Matches:	561
cal Similarity:	99.12%	Conservative:	3
atch:	98.59%	Mismatches:	5
	98.78%	Indels:	0
	21	Gaps:	0

01-168B-2 (1-569) x AAA73131 (1-1707)

```

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
1 ATGGACCCGGCAGAGCGGTGCTGCAGGAGAAGCGCTTAAGTTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
61 AGAGAAGACTGTATATATCGGAAACCCCTAGGAAGATTAATCCAGAGAGAAATTCACCT 120
41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACTTACAACAGCTGTGCCAGCTTTGCATAAACCAAGAGACAGATGTCTAACA 180
61 SerThAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGTATGAAGACTGAAATTTGTGTGGCCAAAGCCAAACTTGCCAAATGGCACTCC 240

```

```

Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db 241 AGCATGATTGTGCCCAAGCAGCGGAAACTCTCAGCAGCTATGAGAAGGAAAGGAGCTG 300
Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 301 TGTGTCAAGTATTATTTGAGCAGTGGTCAGAGTCTGATCAAGTGGAAATTTCTAGAACACCTT 360
Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
Db 361 ATATCCAAATGTCTACCTACAGCATGGGCACATCACTCTCTACCTAAACCTATGCTG 420
Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
Db 421 CAGAGGATTTTCATAACTGCACCTGCCAGCAGCGGGTCTGACCACATCGCTGAGAACATT 480
Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
Db 481 CTGTCTACTTGGACGCCCAAGTCACTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Qy 181 ArgValThrSerAspGlyMetLeuTyrLysLysLeuIleGluArgMetValArgThrAsp 200
Db 541 CGGTGACGTCCGACCGGCATGTGTGGAAGAAAGCTCATCGAGAGGATGGTTCAGGACGAC 600
Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
Db 601 TCTCTGTGCGAGCGCTGCCAGAGCGCAGAGGCTGGGACAGTACTTATTCAAAAACAAA 660
Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
Db 661 CTCCTCTGATGAGAAACGCTCTCCCAACTCTTTATAGAGCGCTTTATCTAAATCATA 720
Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db 721 CAAGACATTGAGACATAGATCCAAATTCGAGATGTGGGGACATAGATTTACAGAGAATC 780
Qy 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
Db 781 CACTGCCGAGTGAACAAAGTAAAGGGGTTTACTGTTTACAGTACGACGCCAGCAGAGATA 840
Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
Db 841 GTCAGCGGCTTCGACAGCAACACCATCAAGATCTGGATAAAGACACACTGGAAATGCAAG 900
Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 901 CGGATTCTCAGGGGCCACACGCGCTCCGTCTGTGTCTGCAGTACGATGAGAGGGTGTATC 960
Qy 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
Db 961 ATCAGAGGCTCTCTCAGACTCCACCGTCAGAGTGTGGGATGTAATGCAGGTGAGATGCTA 1020
Qy 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db 1021 AACACATTGATTCAACCACTGTGAAGCGGTTCTGCACCTCGCTTCAATTAATGGCATGATG 1080
Qy 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
Db 1081 GTACCTGTTTCCAAAGACCGCTTCCATCGCTGTGTGGGATATGGCTTCCCCAACTGACATC 1140
Qy 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
Db 1141 ACCCTCAGAGGGGTGTGTGTGGACACCGAGCTGCGGTCAATGTTGTAGACTTTGATGAC 1200
Qy 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
Db 1201 AAGTACATCGTTTCTCGCTCTGGAGATAGAACATAAAGGTGTGGAACACAAAGTACCTGT 1260
Qy 421 GlyPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
Db 1261 GAATTCGTAGGACCTTAATGGCAACAGCGTGGCATCGCTGTTTGCAGTACAGAGAC 1320
Qy 441 ArgLeuValValSerGlySerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460

```

1321 AGCGTGTGTGAGCGGCTCTCTGACACACCATCAGCTGTGGACATAGTGTGGA 1380
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1381 GCATGCTGCGAGTGTGGAGGCCATGAGGAGTGTGACGTGCTATTCGATTGATTAAC 1440
 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTyrAspLeuValAla 500
 1441 AAAAGGATAGTAGCGGAGCTATGATGGGAAATTAAGTGTGGATCTTATGCGTCT 1500
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1501 TTGGACCCGCGTCTCCAGCAGGAGATCTCTGTCTCGGACACTTGTGGAGCATCTCGA 1560
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1561 AGAGTTTTCGCGCTCCAGTTTGATGAATCCAGATTGTCCAGTAGTTCACATGATGACACA 1620
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1621 ATTCTCATCTGGGACTTCTCGAATGATCCAGTGTCTACGCTGAACCGCGCTCCCT 1680
 561 SerArgThrTyrThrTyrIleSerArg 569
 1681 TCTCGGACATACACTTACTCTCCAGA 1707

13

4

293714 standard; DNA; 2175 BP.

293714;

-AUG-2000 (first entry)

box protein FWDip coding sequence.

ubiquitin ligase; SCF; F-box protein; targeted degradation;
 stabilisation; proteolysis; drug discovery; gene therapy; cancer;
 protein; Huntington's disease; gene knockout; delivery systems;
 ser; SS.

; musculus.

Location/Qualifiers

91..1800

/*tag-a

/product= F-box protein FWDip

00022110-A2.

APR-2000.

OCT-1999; 99WO-US23705.

OCT-1998; 98US-0103787.

RD | HARVARD COLLEGE.

u P, Howley P;

; 2000-317970/27.

SDB; AAY83254.

getting degradation of polypeptide useful for treating cancer and
 er proliferative disorders, involves conjugating polypeptide with
 quitin protein ligase or inhibiting ubiquitination using organic
 bound

im 10; Page 184; 18spp; English.

F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 ases) which can be used for the targeted degradation of a target

CC polypeptide in vivo. Targetted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or
 CC increasing the level of a target polypeptide and for creating and
 CC expressing a destabilized polypeptide which is subjected to SCF
 CC mediated proteolysis. Degrading any desired protein in a cell is
 CC useful for preventing or treating diseases caused by the presence of
 CC abnormal amount of the specific polypeptides, for drug discovery and
 CC for gene therapy. Diseases treated include cancer, by degradation of
 CC oncoproteins, Huntington's disease, other proliferative disorders and
 CC microbial infections. The method provides a quick and easy
 CC alternative to gene knockout technology. The target polypeptide can
 CC be degraded at all stages, or a specific stage, of development in the
 CC mature animal.

SQ Sequence 2175 BP; 589 A; 528 C; 567 G; 491 T; 0 other;

Alignment Scores:

Pred. No.: 3,83e-299 Length: 2175
 Score: 2997.00 Matches: 561
 Percent Similarity: 99.12% Conservative: 3
 Best Local Similarity: 98.59% Mismatches: 5
 Query Match: 98.78% Indels: 0
 DB: 21 Gaps: 0

US-09-601-168B-2 (1-569) x AA293714 (1-2175)

QY 1 MetAspProAlaGluAlaValLeuGlnGlyLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 91 ATGGACCCGCGAGCGGTGCTGCAGGAGAAAGCGCTTAAGTTTATGAATTCCTCAGAG 150
 QY 21 ArgGluAspCysAsnAsnGlyGluProProArgGlyIleIleProGluLysAsnSerLeu 40
 Db 151 AGAGAGACTGTAAATAATGGCGAACCCCTAGAGAGATAATACCAGAGAAAGAAATTCATT 210
 QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 211 AGACAGACTTACACAGCTGTGCCAGGCTTTCATATAAACCAAGACAGATGTCTAACAA 270
 QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 271 AGCACTGCTATGAAGACTGAAAATTTGTGTGGCCAAAGCCAAACTTGCCTAATGGCACTTCC 330
 QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 331 AGCATGATTTGTGCCCAAGCAGCGGAAACTCTCAGCAAGCTATGAGAGAGAAAGAGGCTG 390
 QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 391 TGTGTCAAGTATTTTGTGAGCAGTGGTCAGAGTCTGATCAAGTGGAAATTTGTAGAACACCTT 450
 QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 451 ATATCCCAATGTGTCTACCTACAGCATGGGCGACATCAACTCTCTACTAAACCTATGCTG 510
 QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 511 CAGAGGGATTTCACTAATCTGCACTGCCAGCAGCGGGTCTGGACCATCGCTGAGAACATT 570
 QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 571 CTGTCTACTTGGACGCCAAGTCACTGTGTGTCTGTCTGAGCTCGTGTGCAAGGAATGTGTAC 630
 QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 631 CGGCTGACGTCGGACGGGCTGTCTGTGGAAAAGCTCATCGAGAGGATGCTCAGGACGGAC 690
 QY 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
 Db 691 TCTCTGTGGCAGGCTGTGGCAGGCGCAGAGGCTGGGACAGTACTTATTTCAAAAAACAAA 750
 QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240

```

|||||
751 CCTCTGATGAGAACGCTCTCCCACTCTTTTATAGAGCGTTTATCCTTAATCATCA 810
AC
XX
AAH89966;
XX
01-OCT-2001 (first entry)
XX
Human bone marrow cDNA, SEQ ID NO: 97.
DE
XX
Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
XX
antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
XX
immunosuppressive; gene therapy; cytokine cell proliferation;
XX
cell differentiation modulator; immune disorder; infection; cancer;
XX
human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
OS
Homo sapiens.
XX
WO200153453-A2.
XX
26-JUL-2001.
XX
23-DEC-2000; 2000WO-US34960.
XX
21-JAN-2000; 2000US-0488725.
XX
25-APR-2000; 2000US-0552317.
XX
09-JUL-2000; 2000US-0598042.
XX
19-JUL-2000; 2000US-0620312.
XX
03-AUG-2000; 2000US-0653450.
XX
14-SEP-2000; 2000US-0662191.
XX
19-OCT-2000; 2000US-0693036.
XX
30-NOV-2000; 2000US-0250583.
XX
(HYSE-) HYSEQ INC.
XX
Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX
Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX
Zhao QA, Zhou P, Drmanac RT;
XX
WFI; 2001-488707/53.
XX
P-PSDB; AAM00847.
XX
Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX
for treating e.g. cancer and immune deficiency disorders -
XX
Claim 1; Page 274-275; 648pp; English.
XX
The present sequence is one of 251 novel human polynucleotides
XX
expressed in the bone marrow. The polynucleotide and the
XX
polypeptide encoded by it are useful in the treatment of various
XX
immune deficiencies and disorders. The deficiencies and disorders may
XX
be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
XX
infection, or may result from an autoimmune disorder; a coagulation
XX
disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX
suppression of an inflammatory response or treatment of a nervous
XX
system disorder such as Alzheimer's disease. Detection of the presence
XX
or increased expression of the polynucleotide or the protein it
XX
encodes is useful for the diagnosis and/or prognosis of one
XX
or more types of cancer. The polynucleotide and polypeptide can be
XX
used as nutritional sources or supplements and in the screening of
XX
chemical compounds as potential drugs.
XX
SQ Sequence 3622 BP; 961 A; 839 C; 874 G; 948 T; 0 other;

Alignment Scores:
Pred. No.: 2,12e-298 Length: 3622
Score: 2993.00 Matches: 569
Percent Similarity: 93.89% Conservative: 0
Best Local Similarity: 93.89% Mismatches: 0
Query Match: 98.65% Indels: 37
DB: 22 Gaps: 1

US-09-601-168B-2 (1-569) x AAH89966 (1-3622)
Cy 1 MetAspProAlaGluAlaValLeuGlnGluIysAla-LeuIysPheMet----- 16

```

```

110 ATGGACCGCGCGAGCGGTGCTGCAAGAGAGGACCTCAAGTTTATGTGCTCTATGCC 169
16 ----- 16
170 CAGGTCTCTGTGGCTGGGCTGCTCAGCGCTGGCGGACAGCATGCTTCGCTGGCATGCCT 229
17 -----AenSerSerGluArgGluAspCy 24
230 GTATAACCCAGGCACTGGCGCACTCACAGCTTTCAGAAATTCCTCAGAGAGAAGACTG 289
24 sAsnAsnGlyGluProProArgLysIleIleProGluLysSerSerLeuArgGlnThrTy 44
290 TAATAATGCGAACCCCTTAGAAGATTAATACAGAGAAGAAATTCACCTTAGACAGACATA 349
44 rAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMe 64
350 CAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTTAGCAAGCACTGCTAT 409
64 tLysThrGluAenCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVa 84
410 GAAGACTGAGAAATGTGTGGCCAAACAAACCTTGCCAAATGGCACTTCACGATGATGT 469
84 lProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCysValLysTy 104
470 GCCCAAGCAACGGAACTCTCAGCAAGCTATGAAAGGAAGGAAGCACTGTGTGTCAAATA 529
104 rPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSerGlnMe 124
530 CTTTGAGCAGTGTGTCAGAGTCAGATCAAGTGAATTTGTGGAACATCTTATATCCCAAT 589
124 tCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPh 144
590 GTGTCTTACCACATGGGCATTAACCTGATCTTAAACCTATGTTGGAGAGAGATT 649
144 elleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAenIleLeuSerTyrLe 164
650 CATAACTGCTCTGCCAGCTCGGGATGGATCATATCGCTGAGAAACATTCGTGCATACCT 709
164 uAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSe 184
710 GGATGCCAAATCACTATGTGCTGCTGAACTTGTGCAAGGAATGGTACCGAGTGACCTC 769
184 rAspGlyMetLeuTrpLysLeuIleGluArgMetValArgThrAspSerLeuTrpAr 204
770 TGATGGCATGCTGTGAAGAGCTTATCAGAGAATGGTCAGACAGATTCTCTGTGGAG 829
204 gGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGl 224
830 AGGCTTGGCAGAACGAGAGGATGGGACAGATATTATTCAAAAAACAAACCTCTCTGACGG 889
224 yAenAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGl 244
890 GAATGCTCTCCCACTCTTTTATAGACATTTATCTCTAAATATACAGACATTGA 949
244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
950 GACAAATAGAAATCTAAATTGGAGATGTGGAAGACATAGTTTACAGAGAATTCACGTGCCGAAG 1009
264 rGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLe 284
1010 TGAACAAGCAAGCAAGAGGATTTACTGTTTACAGATATGATATGCAAAATAGTAAGCGGCT 1069
284 uArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuTh 304
1070 TCAGACACACACAAATCAAGATCTGGGATTAATAACACATTTGGATGCAAGCGAATTCCTAC 1129
304 rGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySe 324
1130 AGGCCATACAGGTTACGCTCTGCTCTCAGTATGATGAGAGAGTGATCATAACAGGATC 1189
324 rSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIl 1344

```

RESULT 15

```

AAK51717
ID AAK51717 standard; cDNA; 2366 BP.
XX AAK51717;
XX
XX 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 262.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation; ss.
OS Homo sapiens.
XX
XX W0200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX

```

FEB-2000; 2000US-0496914.
 APR-2000; 2000US-0560875.
 JUN-2000; 2000US-0598075.
 JUL-2000; 2000US-0620325.
 SEP-2000; 2000US-0654936.
 SEP-2000; 2000US-0663561.
 OCT-2000; 2000US-0693325.
 NOV-2000; 2000US-0728422.

(XSE-) HYSEQ INC.

g YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 e AU, Yang Y, Wejhrman T, Goodrich R;

; 2001-476283/51.

SDB; AAK78584.

leic acids encoding polypeptides with cytokine-like activities,
 ful in diagnosis and gene therapy -

im 1; Page 1183-1186; 6221pp; English.

invention relates to polynucleotides (AAK51456-AAK53435) and the
 oded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 okine, cell proliferation or cell differentiation or which may induce
 uction of other cytokines in other cell populations. The
 ynucleotides and polypeptides are useful in gene therapy, vaccines or
 ide therapy. The polypeptides have various cytokine-like activities,
 stem cell growth factor activity, haematopoiesis regulating
 ivity, tissue growth factor activity, immunomodulatory activity and
 ivan/inhibin activity and may be useful in the diagnosis and/or
 eatment of cancer, leukaemia, nervous system disorders, arthritis and
 lamation.

e: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 MM80020) are omitted as the relevant pages from the sequence listing
 e missing at the time of publication.

quence 2366 BP; 661 A; 519 C; 594 G; 592 T; 0 other;

t Scores:

1.27e-298 Length: 2366
 2992.50 Matches: 569
 Similarity: 90.03%
 al Similarity: 90.03%
 Mismatches: 0
 Indels: 63
 Gaps: 1

1-168B-2 (1-569) x AAK51717 (1-2366)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLys----- 14
 127 ATGACCGCGCGAGGGGCTCTGCAAGAGAAGGCACTCAAGTTTATGATGGAGTTTCG 186
 14 ----- 14
 187 TCTTGGTCCAGCGTGAATACAATGGCGCGATCTCGGCTACCGCAACCTCCACCTCC 246
 14 ----- 14
 247 CGGGTTCAATGCTCTATGCCAGGCTCTGTGGTGGGCTGCTCCAGCCTGGCGGACGC 306
 15 -----PheMetAsn 17
 307 ATGCTTCGCTCGCATGCTGTATATACCCAGGAGCTGGCGGCACTCACAGCTTTTCATGA 366
 18 SerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLys 37
 367 TCCTCAGAGAGAAGACTGTAATAATGGCGAACCCTTAGAAGATAATACCAGAGAAG 426
 38 AsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrVal 57
 427 AATTCGCTTAGACAGACATACAACAGCTGTCCAGACTCTCTGCTTTAAACCAAGAAACAGTA 486

QY 58 CysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsn 77
 |||||
 Db 487 TGTTTAGCAAGCACTGCTATGAAGACTGAGAATTCGTGGCCAAACAAACCTTCCCAAT 546
 |||||
 QY 78 GlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGlu 97
 |||||
 Db 547 GGCACCTTCAGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAA 606
 |||||
 QY 98 LysGluLeuCysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheVal 117
 |||||
 Db 607 AAGGAACCTGTGTGTCARAATACCTTTGAGCAGTGGTCCAGAGTCAGATCAAGTGAATTTGTG 666
 |||||
 QY 118 GluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLys 137
 |||||
 Db 667 GAACATCTATATCCAAATGTCTATTACCAACATGGGCACATAAACTCGTATCTTAA 726
 |||||
 QY 138 ProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAla 157
 |||||
 Db 727 CCTATTGTGCAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGGATTGGATCATATTGT 786
 |||||
 QY 158 GluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLys 177
 |||||
 Db 787 GAGAACATTTCTGTATACCTGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTGCAAG 846
 |||||
 QY 178 GluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetVal 197
 |||||
 Db 847 GAATGGTACCGAGTGACCTCTGATGCGATGCTGTGGAAGAGCTTATCGAGAGAATGGTC 906
 |||||
 QY 198 ArgThrAspSerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPhe 217
 |||||
 Db 907 AGCAGAGATTCTCTGTGGAGAGCGCTGCGAGAACGAGAGGATGGGAGCATGATTATTTC 966
 |||||
 QY 218 LysAsnLysProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrPro 237
 |||||
 Db 967 AAAACAACACTCTCTGACGGGAATGCTCTCCCACTCTTTTATAGAGCACTTTATCT 1026
 |||||
 QY 238 LysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeu 257
 |||||
 Db 1027 AAAATTATACAAGACATTGAGACAATAGAATCTAATTGGAGATGTGGAAGACATAGTTTA 1086
 |||||
 QY 258 GlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAsp 277
 |||||
 Db 1087 CAGAGAAATTCACCTGCCGAAGTGAACAAGCAAGAGGAGTTTACTGTTTACAGTATGATGAT 1146
 |||||
 QY 278 GlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeu 297
 |||||
 Db 1147 CAGAAAATAGTAAGCGGCTTCGAGACAACACAAATCAAGATCTGGGATAAAAACACATTG 1206
 |||||
 QY 298 GluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu 317
 |||||
 Db 1207 GAATGCAAGCGAAATCTCACAGGCCATACAGGTTTCAGTCTCTGTCTCCAGTATGATGAG 1266
 |||||
 QY 318 ArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGly 337
 |||||
 Db 1267 AGAGTGATCATAAACAGGATCATCGGATTCACGGTCAGAGTGTGGGATGTAATAACAGGT 1326
 |||||
 QY 338 GluMetLeuAsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsn 357
 |||||
 Db 1327 GAAATGCTAAACACGTTGATTCACCATTTGTGAAGCAGTCTGCACCTTCGGTTTCAATAAT 1386
 |||||
 QY 358 GlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerPro 377
 |||||
 Db 1387 GGCATGATGGTGACCTCTCCAAAGATCGTTCATTCGTGTATGGGATATGGCCTCCCA 1446
 |||||
 QY 378 ThrAspIleThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAsp 397
 |||||
 Db 1447 ACTGACATTACCCTCCGAGGGTGTGTGCGACACCGAGCTGCTGTCATTTGTGTAGAC 1506
 |||||
 QY 398 PheAspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThr 417
 |||||
 Db 1507 TTTTGATGACAGTACATTGTTTCTGCATCTCGGGGATAGAACTATTAAGGTATGGAAACACA 1566
 |||||

418 SerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGln 437
|||||
1567 AGTACTTGTGAATTCTGAAGGACCTTAAATGGACACAAACGAGGCATTGCCTGTTTGAG 1626
|||||
438 TyrArgAspArgLeuValSerGlySerSerAspAsnThrIleArgLeuTrpAspIle 457
|||||
1627 TACAGGACAGGCTGGTGTGTGCTCATCTTGACACACTATCAGATTATGGGACATA 1686
|||||
458 GluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArg 477
|||||
1687 GAATGTGGTGCATGTTTACGAGTGTAGAGGCCATGAGCAATTGGTGGTGTATTGGA 1746
|||||
478 PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeu 497
|||||
1747 TTTGATTAACAAGAGGATAGTCAGTGGGGCTATGATGGAAATTAAGTGTGGGATCTT 1806
|||||
498 ValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGlu 517
|||||
1807 GTGGCTGCTTTGGACCCCGTCTCTCGCAGGGACACTCTGTCTACGGACCCCTTGTGGAG 1866
|||||
518 HisSerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHis 537
|||||
1867 CATTCGGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCCAGATTGCAGATTACAT 1926
|||||
538 AspAspThrIleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProPro 557
|||||
1927 GATGACACAAATCCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCC 1986
|||||
558 ArgSerProSerArgThrTyrThrTyrIleSerArg 569
|||||
1987 CGTTCCCTTCTCGAACATACACCTACATCTCCAGA 2022
|||||

Completed: October 22, 2003, 15:03:20

S : 422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

tein - nucleic search, using frame_plus_p2n model

October 22, 2003, 16:26:53 : Search time 395 Seconds
(without alignments)
3963.493 Million cell updates/sec

US-09-601-168B-2
score: 3034
1 MDPRAVLQKALKEFNSE.....PAAQAEPRSPSRITYVISR 569

table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 1792395 seqs, 1340900451 residues

umber of hits satisfying chosen parameters: 3584790

DB seq length: 0

DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

line parameters:

frame_p2n.model -DEV-xlp
!_1/USFTO.spool_p/US09601168/runat_22102003.100548.17891/app_query.fasta.1.711
ished Applications NA -QRMFASTAP -SUFFIX=rnpb -MINMATCH=0.1
-0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=BLOSUM62
uman40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
v=0 -ALIGN=15 -MODS=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
2000000000 -USER=US09601168 @CNG.1.1.95 @runat_22102003.100548.17891
-ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTVS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

ed. No. is the number of results predicted by chance to have a
ore greater than or equal to the score of the result being printed,
id is derived by analysis of the total score distribution.

SUMMARIES

Query	Score	Match	Length	ID	Description
3034	100.0	1710	14	US-10-038-010-7	Sequence 7, Appli

2	3034	100.0	2151	12	US-09-832-161-17	Sequence 17, Appli
3	3034	100.0	2151	13	US-10-042-417-1	Sequence 1, Appli
4	2384.5	78.6	4230	12	US-09-832-161-15	Sequence 15, Appli
5	1163	38.3	657	14	US-10-023-530-1	Sequence 1, Appli
6	1155.5	38.1	951	9	US-09-764-848-15	Sequence 15, Appli
7	1155.5	38.1	951	12	US-10-222-020-15	Sequence 15, Appli
8	1155.5	38.1	951	14	US-10-116-016-15	Sequence 15, Appli
9	689	22.7	479	9	US-09-864-781-2778	Sequence 2778, Ap
10	644.5	21.2	1881	10	US-09-213-888-20	Sequence 20, Appli
11	644.5	21.2	1881	10	US-09-328-877A-13	Sequence 20, Appli
12	640	21.1	1620	12	US-10-245-618-20	Sequence 13, Appli
13	640	21.1	1659	12	US-10-245-618-7	Sequence 7, Appli
14	640	21.1	1770	12	US-10-245-618-17	Sequence 17, Appli
15	640	21.1	1884	12	US-10-245-618-5	Sequence 5, Appli
16	640	21.1	2001	10	US-09-213-888-26	Sequence 26, Appli
17	640	21.1	2001	10	US-09-328-877A-26	Sequence 26, Appli
18	640	21.1	2010	10	US-09-213-888-24	Sequence 24, Appli
19	640	21.1	2010	10	US-09-328-877A-24	Sequence 24, Appli
20	640	21.1	2063	12	US-10-245-618-34	Sequence 34, Appli
21	640	21.1	2124	12	US-10-245-618-1	Sequence 1, Appli
22	640	21.1	2255	12	US-10-245-618-35	Sequence 35, Appli
23	640	21.1	3550	10	US-09-213-888-1	Sequence 1, Appli
24	640	21.1	3571	10	US-09-328-877A-1	Sequence 1, Appli
25	640	21.1	3571	10	US-09-213-888-2	Sequence 2, Appli
26	640	21.1	3571	10	US-09-328-877A-2	Sequence 2, Appli
27	575	19.0	2037	10	US-09-801-368-313	Sequence 313, App
28	552.5	18.2	2091	14	US-10-128-714-2506	Sequence 2506, Ap
29	552.5	18.2	2091	14	US-10-128-714-7506	Sequence 7506, Ap
30	552.5	18.2	2167	14	US-10-128-714-1506	Sequence 1506, Ap
31	552.5	18.2	2167	14	US-10-128-714-6506	Sequence 6506, Ap
32	552.5	18.2	4167	14	US-10-128-714-506	Sequence 506, App
33	552.5	18.2	4167	14	US-10-128-714-5506	Sequence 5506, App
34	519.5	17.1	2208	9	US-10-032-585-6658	Sequence 6658, Ap
35	517	17.0	370	9	US-09-864-761-19492	Sequence 19492, A
36	414	13.6	524	12	US-10-029-386-4953	Sequence 4953, Ap
37	405	13.3	250	12	US-10-029-386-18709	Sequence 18709, A
38	392	12.9	2237	10	US-09-994-485-7	Sequence 7, Appli
39	392	12.9	2237	10	US-09-832-232-11	Sequence 11, Appli
40	344	11.3	1548	14	US-10-128-714-7213	Sequence 7213, Ap
41	342	11.3	319	10	US-09-960-352-2934	Sequence 2934, Ap
42	338.5	11.2	1923	14	US-10-102-806-190	Sequence 190, App
43	338.5	11.2	3747	14	US-10-197-666A-143	Sequence 143, App
44	338.5	11.2	3864	14	US-10-197-666A-149	Sequence 149, App
45	337	11.1	3465	10	US-09-994-485-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-10-038-010-7
; Sequence 7, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2002-07-23
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1710)
; OTHER INFORMATION: Human beta-TrCP1 : F-box containing protein with 7 WD40 repeats;
; OTHER INFORMATION: Part of SCF (Skp1/Cullin/F-box) complex E3 ubiquitin ligase; Impl
; OTHER INFORMATION: located in the degradation of b-catenin and Ikba

38-010-7

nt Scores:

0 Length: 1710
Matches: 569
Similarity: 3034.00
Conservative: 0
Mismatch: 100.00%
Indels: 0
Gaps: 0
atc: 14

31-168B-2 (1-569) x US-10-038-010-7 (1-1710)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
1 ATGGACCCGGCGAGCGGTGCTGCAAGAGAGAGGACCTCAAGTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
61 AGAGAGAGACTGTAATAATGCGCAACCCCTAGGAAGATAATACCAAGAGAAATTCACCT 120
41 ArgGlnThrTyrrAsnSerCysAlaArgLeuLysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACATACACAGCTGTGCCAGACTCTGCTTTAAACCAGAAACAGTATGTTTAGCA 180
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGCTATCAAGACTGAGAATTTGTGTGGCAAAACAAACATGGCAATGGCACTTCC 240
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLysGluLysGluLeu 100
241 AGTATGATTGTGCCCAAGCAAGCAAGCAATCTCAGCAAGCTATGAAAGAAAGAAAGACTG 300
101 CysValLysTyrrPheGluGlnThrPheGluSerAspGlnValGluPheValGluHisLeu 120
301 TGTGTCAAAATCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGAATTTGTGGAACATCTT 360
121 IleSerGlnMetCysHisTyrrGlnHisGlyHisIleAsnSerTyrrLeuLysProMetLeu 140
361 ATATCCAAATGTGTCATACCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAAT 420
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
421 CAGAGAGATTCATACACTCTCTGCCAGCTCGGAGATGGATCATATCGCTGAGAACAT 480
161 LeuSerTyrrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGlnTyrr 180
481 CTGTCTACTCTGGATGCCAAATCACTATGCTGCTGAACCTTGTGTGCAAGAAATGGTAC 540
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
541 CGAGTGACCTCTGATGGCATGCTGTGGAGAGAGCTTATCGAGAGAAATGGTCAGGACAGAT 600
201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrrLeuPheLysAsnLys 220
601 TCTCTGTGGAGAGCGCTGGCAGAACAGAGAGATGGGACAGATATTTATTCAAAACAAA 660
221 ProProAspGlyAsnAlaProProAsnSerPheTyrrArgAlaLeuTyrrProLysIleLeu 240
661 CCTCTCTGAGCGGAGATGCTCTCCCACTCTTTTATAGGACATTTATCTTAAATATATA 720
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
721 CAAGACATTTGAGACATAGAAATCTAATTTGGAGATGGGAGACATAGTTTACAGAGAAAT 780
261 HisCysArgSerGluThrSerLysGlyValTyrrCysLeuGlnTyrrAspAspGlnLysIle 280
781 CACTGCGGAGTGAACCAAGCAAGAGATTTACTGTTTACAGTATGATGATCAGAAATA 840
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
841 GTAAGCGGCTTCGACACACACAACTCAGATCTGGGATAAACACACATTTGAATGCAAG 900
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrrAspGluArgValIle 320

Db 901 CGAATTCCTCAGAGCCATACAGGTTCAAGTCTCTGTCTCCAGTATGATGAGAGAGTGATC 960
QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 961 ATACAGAGATCATCGGATTCACCGGTGAGAGTGGGATGTAATACAGGTGAAATGCTA 1020
QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db 1021 AACACGTTGATTCACCATTTGAAGCAGCTTCGACATTCGCTTTCATATAATGGCATGATG 1080
QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
Db 1081 GTGACCTGCTCAAAAGATCGTTCATTTGCTATGAGGATATGGCTCCCAACTGACATT 1140
QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
Db 1141 ACCCTCCGAGGGTGTGCTCGGACACCGAGCTGCTCAATGTTGTAGACTTTTGATGAC 1200
QY 401 LysTyrrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
Db 1201 AGATACATTTGTTCTGCACTCTGGGATAGAACTATAAGGTATGGAACACAGTACTTGT 1260
QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrrArgAsp 440
Db 1261 GAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTGCAGTACAGGGAC 1320
QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
Db 1321 AGCTGTGTAGTGTGGTCTCATCTGCACACACTATCAGATTATGGGACATAGATGTGGT 1380
QY 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysLysIleArgPheAspAsn 480
Db 1381 GCATGTTTACAGAGTGTAGAGGCAATGAGGAAATTTGGTGCCTTGTATTCGATTTGATAAC 1440
QY 481 LysArgIleValSerGlyValTyrrAspGlyLysLysValTrpAspLeuValAlaAla 500
Db 1441 AAGAGATAGTCAAGTGGGGCTTATGATGGAATAATTAAGTGGGATCTTTGGCTGCT 1500
QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
Db 1501 TTGGACCCCGTGTCTCTGCAGGAGACACTGTCTACGGACCTTTGTGGAGCATTCGGGA 1560
QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540
Db 1561 AAGATTTTTCGACTACAGTTTGATGAATTCAGATTTGTCAGATTTCCATGATGACACA 1620
QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
Db 1621 ATCTCATCTGGGACTTCTTAAATGATCCAGCTGCCAAGCTGAACCCCGCTTCCCT 1680
QY 561 SerArgThrTyrrThrTyrrIleSerArg 569
Db 1681 TCTCGAACATACACCTACATCTCCAGA 1707

RESULT 2

US-09-832-161-17
; Sequence 17, Application US/09832161
; Publication No. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercurio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinon
; APPLICANT: Davis, Matti
; APPLICANT: Hatzuba, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: 860098.427
; CURRENT APPLICATION NUMBER: US/09/832,161
; CURRENT FILING DATE: 2001-04-09

APPLICATION NUMBER: 09/210,060

FILING DATE: 1998-12-10

R OF SEQ ID NOS: 30

ARE: PatentIn Ver. 2.0

NO 17

TH: 2151

: DNA

NISM: Homo sapiens

2-161-17

t Scores:

.: 0 Length: 2151

Matches: 569

Similarity: 3034.00

Conservative: 0

Mismatch: 100.00%

Indels: 0

Gaps: 0

1-168B-2 (1-569) x US-09-832-161-17 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20

70 ATGACCCCGCGGAGGGGCTGCTCCAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129

21 ArgGluAspCysAsnAsnGlyGluProProArgLysLeileProGluLysAsnSerLeu 40

130 AGAAGAGACTGTAATAATGGCGAACCCCTAGGAAGATAATACCAGAGAGAAGATTCACTT 189

41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60

190 AGACAGACATACAAACAGCTGTCCTTTAAACCAAGAAACAGATGTTTAGCA 249

61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80

250 AGCACTGCTATGAGACTGAGATTGTGGCCAAACAAACTTGGCAATGGCACTTCC 309

81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100

310 AGTATGATTGTGCCCAAGCAAGAACTCTCAGCAAGCTATGAAAAGGAAAGAACTG 369

101 CysValLysTyrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120

370 TGTGTCAAATCTTTGAGCAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTGGAACTCTT 429

121 IleSerGlnMetCysHisTyrGlnHisLysIleAsnSerTyrLeuLysProMetLeu 140

430 ATATCCCAATGCTCATACCAACATGGGCACATAAAGCTGATCTTTAAACCTATGTTG 489

141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160

490 CAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGGATTGGATCATATCGCTGAGAACATT 549

161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluThrTyr 180

550 CTGTCACTACCTGGATGCCAATCACTATGCTCTGCAACTTGTGTCCAAAGGAATGGTAC 609

181 ArgValThrSerAspGlyMetLeuTyrLysLysLeuIleGluArgMetValArgThrAsp 200

610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAAATGGTCAGACAGAT 669

201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220

670 TCTCTGTGGAGAGGCTGGCAGAACGAAGAGATGGGAGATGTTATTCAAAACAAA 729

221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240

730 CCTCCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACCTTTATCTCTAAATATATA 789

241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260

790 CAAACATTTGACACATAGAAATCTAATTGGAGATGTGGAACACATAGTTTACAGAGAATT 849

261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280

RESULT 3

US-10-042-417-1

; Sequence 1, Application US/10042417

; Publication No. US20020123082A1

; GENERAL INFORMATION:

; APPLICANT: Pagano, M.

; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF

; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS

; FILE REFERENCE: 5914-090-999

; CURRENT APPLICATION NUMBER: US/10/042,417

Db 850 CACTGCCGAGTGAACAAGCAAGAGTTTACTGTTTACAGTATGATCAGAAATA 909

QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300

Db 910 GTAAGCGGCTTCGAGACAACAACATCAAGATCTGGGATAAAACACACATTGGAATGCAAG 969

QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320

Db 970 CGAATTCTCAGAGGCCATACAGTTTCAGTCTCTGTTCCAGTATGATGAGAGATGATC 1029

QY 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340

Db 1030 ATAACAGGATCATCGGATTCACGCTGAGAGTGGGATGTAATACAGTGAATGCTA 1089

QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360

Db 1090 AACACGTTGATTCACCAATTGTGAAGCAGTTCTGCACCTTCGCTTCAATAATGSCATGATG 1149

QY 361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380

Db 1150 GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCTCCCAACTGACATT 1209

QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400

Db 1210 ACCCTCCGAGGGTGTGTCGACACCGAGCTGCTGTCATGTTGTAGACTTTGATGAC 1269

QY 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420

Db 1270 AAGTACATTGTTTCTGCACTCTCGGGATAGAACTATAAAGGTATGGAACACAAGTACTTGT 1329

QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440

Db 1330 GAATTGTAAAGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTGCAGTACAGGAC 1389

QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460

Db 1390 AGGCTGGTAGTCAGTGGCTCACTGCAACAACACTATCAGATTATGGACATAGAAATGTGCT 1449

QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480

Db 1450 GATGTTTACAGAGTGTTAGAAGGCCATGAGGAATTTGGTGCCTGTTATTCGATTTGATAAC 1509

QY 481 LysArgIleValSerGlyValatTyrAspGlyLysIleLysValTyrAspLeuValAlaAla 500

Db 1510 AAGAGGATAGTCACTGGGGCTTATGATGGAAAAAATTAAGTGTGGATCTTGTGCTGCT 1569

QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLysLeuArgThrLeuValGluHisSerGly 520

Db 1570 TTGGACCCCGTGTCTCGAGGACACTCTGCTACGGACCTTGTGGAGCATTCGGA 1629

QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540

Db 1630 AGAGTTTTCGACTACAGTTTGCATGAATTCAGATTTCAGTAGTTTCACATGATGACACA 1689

QY 541 IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560

Db 1690 ATCCTCATCTGGGACTTCTTAATATGATCCAGCTGCCAAGCTGAACCCCCCTTCCCT 1749

QY 561 SerArgThrTyrThrTyrIleSerArg 569

Db 1750 TCTCGAACATACACCTACATCTCCAGA 1776

NT FILING DATE: 2002-01-07
 APPLICATION NUMBER: 60/260,179
 FILING DATE: 2001-01-5
 R OF SEQ ID NOS: 89
 ARE: PatentIn Ver. 2.0
 NO 1
 TH: 2151
 : DNA
 : USNM: Homo sapiens
 2-417-1

t Scores:

Length: 2151
 Matches: 569
 Similarity: 3034.00
 Conservativity: 0
 al Similarity: 100.00%
 Mismatches: 0
 Indels: 0
 Gaps: 0

1-168B-2 (1-569) x US-10-042-417-1 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 70 ATGGACCGGCGGAGCGGTGCTGCAAGAGAGAGCGCACTCAAGTTTATGAATTCCTCAGAG 129
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
 130 AGAAGAACTGTATAATATGGCAACCCCTAGGAAGATAATACCAAGAGAAATTCACCT 189
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 190 AGACAGACATACACACACTGTGCAGACTCTGCTTAAACCAAGAACAGATATGTTAGCA 249
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 250 AGCACTGCTATGAAGACTGAGAAATGTGTGGCCAAAACAAACTTGCCAATGGCACTTCC 309
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 310 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAAAGGAACCTG 369
 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 370 TGTGTCAATACITTTGAGCAGTGTGACAGTCAGATCAGTGGNAITTTGTGGACATCTT 429
 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 430 ATATCCCAATGTGTCATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 499
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 490 CAGAGATTTTCATACCTGCTCCAGCTCGGGATTTGATCATATCGCTGAGAACATT 549
 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTrp 180
 550 CTGTCACTACCTGGATGCCAAATCACTATGTGCTGCACTGTGTGTGCAAGGAATGCTAC 609
 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCGAGAGAAATGGTCCAGACAGAT 669
 201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
 670 TCTCTGTGGAGAGCGCTGGCAGACGAGAGAGATGGGACAGATATTTATTTCAAAAACAAA 729
 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 730 CCTCTGACGGGAATGCTCCTCCCACTCTTTTATAGACACTTTATCTCTAAATATATA 799
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 790 CAAGACATTGAGACAATAGATCTTAATTGGAGATGTGGAACATAGTATTACAGAGAANT 849

QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 Db 850 CACTCCGGAAGTGAACAAGCAAGAGAGTCTTACTGTTTACAGATATGATCATCAGAAAAATA 909
 QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleIleIleLysAsnThrLeuGluCysLys 300
 Db 910 GTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATAAANAACACATTGGAAATGCAG 969
 QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 Db 970 CGAATTCCTCAGAGGCATACAGGTCAGTCTCTCTCCAGTATGATGAGAGAGTGATC 1029
 QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 Db 1030 ATACAGGATCATCGGATTCACGGTCAGAGTGGGATGTAATATACAGGTGAATGCTA 1089
 QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 Db 1090 AACACGTTGATTCACCATTTGTGAAGCAGTCTCGCACTTCGCTTCAATAATGGCATGATG 1149
 QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 Db 1150 GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCTCCCAACTGACATT 1209
 QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 Db 1210 ACCCTCCGAGGGTCTGCTGGCACCGACTGCTGTCATGTTGTAGACTTTGATGAC 1269
 QY 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 Db 1270 AAGTACATGTTTCTGCACTCTGGGATAGAACTATAAGGTATGGAACACAACTACTTGT 1329
 QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 Db 1330 GAATTTGTAAGGACCTTAAATGGACACAAACGAGGCATTGCTGTTTCAGTACAGGAC 1389
 QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 Db 1390 AGCTGGTAGTGAGTGGCTCATCTGACAACTATCAGATTATGGACATAGAAATGTGT 1449
 QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 Db 1450 GCATGTTTACGAGTGTTAGAAGGCCATCAGGAAATTTGGTGGCTGTATTTCGATTGATAAC 1509
 QY 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
 Db 1510 AAGAGGATAGTCAGTGGGCGCTATGATGGAATAATTAAGTGTGGGATCTTTGTGGCTGCT 1569
 QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 Db 1570 TTGGACCCCGGCTCTCGAGGGACACTCTGTCTACGGACCTTTGTGGAGCATTCGGGA 1629
 QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 Db 1630 AGAGATTTTCGACTACAGTTTATGAAATTCAGATTTGTCAGTAGTTCATGATGATGACACA 1689
 QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 Db 1690 ATCTCTCATCTGGGACTTCTTAATGATCCAGTCCCAAGTGAACCCCGGTTCCGCT 1749
 QY 561 SerArgThrTyrThrTyrIleSerArg 569
 Db 1750 TCTCGAACATACACCTACATCTCCAGA 1776

RESULT 4

US-09-832-161-15
 ; Sequence 15, Application US/09832161
 ; Publication NO. US20030166587A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Manning, Anthony M.
 ; APPLICANT: Mercurio, Frank
 ; APPLICANT: Amit, Sharon
 ; APPLICANT: Ben-Neriah, Yinon

S-10-023-530-1					
Sequence 1, Application US/10023530					
Publication No. US20030007956A1					
GENERAL INFORMATION:					
APPLICANT: LEGRAIN, Pierre					
APPLICANT: BENAROUS, Richard					
APPLICANT: BLOT, Guillaume					
APPLICANT: LASSOT, Irina					
TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP					
FILE REFERENCE: B4717A					
CURRENT APPLICATION NUMBER: US/10/023.530					
PRIOR FILING DATE: 2002-04-22					
PRIOR APPLICATION NUMBER: 60/256,276					
PRIOR FILING DATE: 2000-12-18					
NUMBER OF SEQ ID NOS: 30					
SOFTWARE: PatentIn version 3.1					
SEQ ID NO 1					
LENGTH: 657					
TYPE: DNA					
ORGANISM: Homo sapiens					
FEATURE:					
NAME/KEY: gene					
LOCATION: (1)..(657)					
OTHER INFORMATION: Beta TrCP					
IS-10-023-530-1					
Alignment Scores:					
Pred. No.: 2,71e-137	Length: 657				
Score: 1163.00	Matches: 219				
Percent Similarity: 100.00%	Conservative: 0				
Best Local Similarity: 100.00%	Mismatches: 0				
Query Match: 38.33%	Indels: 0				
DB: 14	Gaps: 0				
US-09-601-168B-2 (1-569) x US-10-023-530-1 (1-657)					
Y 1 MetAspProAlaGluValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20					
b 1 ATGGACCCCGGCGAGCGGTGTCTCAAGAAGAAGGCACCTCAAGTTTATGAATTTCCTCAGAG 60					
Y 21 ArgGluAspCysAsnAsnClyGluProProArgLysIlelleProGluLysAsnSerLeu 40					
b 61 AGAGAAGACTGTAAATAATGGGACCCCCTAGAGAGATAATACCAGAGAAGAAATTCAC TT 120					
Y 41 ArgGlnThrTy rAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60					
b 121 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAACCAAGAACAACAGTAGTTTAGCA 180					
Y 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80					
b 181 AGCACTGCTATGAAGACTGAGAAATTTGTGGCCAAAACAAAACITTGCAATGGCACATCC 240					
Y 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTy rGluLysGluLysGluLeu 100					
b 241 AGTATGATTGTCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAGAAGAAAGCAACTG 300					
Y 101 CysValLysTy rPheGluClnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120					
b 301 TGTGTCAAATATCTTTGAGCAGGTGGTCAGAGTCAGATCAAGTGGAAATTTGTGGAACTCTT 360					
Y 121 IleSerGlnMetCysHisTy rGlnHisGlyHisIleAsnSerTy rLeuLysProMetLeu 140					
b 361 ATATCCCAATGTGTTCATACCAACATGGGCACATPAACCTCGTATCTTAAACCTATGTTG 420					
Y 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160					
b 421 CAGAGAGATTTCAATAACTCTCTGCGCAGCTCGGGGATTTGGATCATATCGCTGAGAACAT 480					
Y 161 LeuSerTy rLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTy r 180					
b 481 CTGTCAATCCTGGATGCCAAATCACTATGTCTGCTGAACCTTTGTGCAAGGAATGGTAC 540					
Y 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200					
DB 541 CGAGTGACCTCTGTATGGCATGCTGTGGAAGAAGCTTATCGAGAGAATGCTCAGGACAGAT 600					
QY 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTy rLeuPheLysAsn 219					
DB 601 TCTCTGTGGAGAGCGCTGGCAGAACGAGAGGATGGGACAGTATTTATTCAAAAAC 657					
RESULT 6					
US-09-764-848-15					
; Sequence 15, Application US/09764848					
; Patent No. US2002007270A1					
; GENERAL INFORMATION:					
; APPLICANT: Rosen et al.					
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies					
; FILE REFERENCE: PTZ08					
; CURRENT APPLICATION NUMBER: US/09/764,848					
; CURRENT FILING DATE: 2001-01-17					
; Prior application data removed - consult PALM or file wrapper					
; NUMBER OF SEQ ID NOS: 53					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 15					
; LENGTH: 951					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
; FEATURE:					
; NAME/KEY: SITE					
; LOCATION: (883)					
; OTHER INFORMATION: n equals a,t,g, or c					
; NAME/KEY: SITE					
; LOCATION: (913)					
; OTHER INFORMATION: n equals a,t,g, or c					
; NAME/KEY: SITE					
; LOCATION: (931)					
; OTHER INFORMATION: n equals a,t,g, or c					
; NAME/KEY: SITE					
; LOCATION: (935)					
; OTHER INFORMATION: n equals a,t,g, or c					
; NAME/KEY: SITE					
; LOCATION: (941)					
; OTHER INFORMATION: n equals a,t,g, or c					
US-09-764-848-15					
Alignment Scores:					
Pred. No.: 4.47e-136	Length: 951				
Score: 1155.50	Matches: 226				
Percent Similarity: 76.97%	Conservative: 28				
Best Local Similarity: 68.48%	Mismatches: 37				
Query Match: 38.09%	Indels: 41				
DB: 9	Gaps: 3				
US-09-601-168B-2 (1-569) x US-09-764-848-15 (1-951)					
QY 15 PheMetAsnSerSerGluArgGluAspCysAsnAsnClyGluProProArgLysIlelle 34					
DB 59 TTCAGAGACACTTCAGTTATGGAAGATCAAAATGAGATGAGTCCCCAAAG----- 109					
QY 35 ProGluLysAsnSerLeuArgGlnThrTy rAsnSerCysAlaArgLeuCysLeuAsnGln 54					
DB 110 -----AAAAATATCTTTTGGCAG----- 127					
QY 55 GluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLys 74					
DB 127 ----- 127					
QY 75 LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTy r 94					
DB 128 ATAAGTAATGGAACATCATCTGTCTCCAGAAAGAGGCCATCAGAAAGGAAACTAT 187					
QY 95 GluLysGluLysGluLeuCysValLysTy rPheGluClnTrpSerGluSerAspGlnVal 114					
DB 188 CAAAAGAAAAGAAAGAACTGTGTGTATTAATATTTTCCAGCAGTGGTCTGAATCAGATCAAGTG 247					
QY 115 GluPheValGluHisLeuLeuSerGlnMetCysHisTy rGlnHisGlyHisIleAsnSer 134					

b 248 GAATTGTGGACATCTTTATTTACGAATGTGTCAATATACGATGACATATTAATCTCT 307
Y 135 TyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAsp 154
b 308 TACCTGAGCCCATGTTGACGGGACCTTTATTACCGCTTTACAGAGCAAGCGTTAGAT 367
Y 155 HisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeu 174
b 368 CACATAGCAGAAACATCTTCTGTACCTGGATGCCAGGTCTCTGTGTGAGCAGAGCTG 427
Y 175 ValCysLysGluTyrTrpArgValThrSerAspGlyMetLeuTyrLysLeuIleGlu 194
b 428 GTATGTAAGAATACGACGGAGTGATCTCAGAAGATGCTTTGGAAGAGCTGATTGAA 487
Y 195 ArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGluArgGlyTyrGlyGln 214
b 488 CGAATGGTACGACATGATCCCTATGGAAGAGACTTTTCAGAAAGAGAGGTGGATCAG 547
Y 215 TyrLeuPheLysAsnLysProAspGlyAsnAlaProProAsnSerPheTyrArgAla 234
b 548 TACCTGTTTAAACACAGACCCACAGATGCG-----CCTCCAAATTCATTTTATAGTCA 601
Y 235 LeuTyrProLysIleIleGluAspIleGluThrIleGluSerAsnTrpArgCysGlyArg 254
b 602 TTATACCCAAAGATTATCCAGATATAGACTATAGACTATAGACTACTGCGGTGTGGACGA 661
Y 255 HisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGln 274
b 662 CACAACCTTGCAGAGATTTCAGTGGCGCTCTGAAATAGTAAGGTGTCTACTGTTTACAG 721
Y 275 TyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLys 294
b 722 TAGCATGTGAAAAAATTATCAGTGGCT-CCAGATAATCTATTAGATATGGATAAA 780
Y 295 AsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGln 314
b 781 ACCWCGCTGGAATGTTGAAAGTGTWACAGACACACAGCTCTGK-CTCTGTCTGCAG 839
Y 315 TyrAspGluArgValIleIleThrClySerSerAspSerThrValArgValTyrAspVal 334
b 840 TAATGATGAKCGGTGATTATKTWACTGGCTCTTCAGATTTCTACGGNAGAGGTGGGATGTG 899
Y 335 AsnThrGlyGluMetLeuAsnThrLeuIle 344
b 900 ACACCGGTGAAAGNCTTAACACATGTATC 929

RESULT 7

S-10-222-020-15

Sequence 15, Application US/10222020

Publication No. US20030175739A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT208C2

CURRENT APPLICATION NUMBER: US/10/222,020

CURRENT FILING DATE: 2002-08-16

PRIOR APPLICATION NUMBER: 10/116,016

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: 09/764,848

PRIOR FILING DATE: 2001-01-17

PRIOR APPLICATION NUMBER: 60/179,065

PRIOR FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: 60/180,628

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: 60/214,886

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 60/217,487

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/225,758

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/220,963

PRIOR FILING DATE: 2000-07-26

PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039

PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,038
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/236,370
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/236,802
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,037
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,040
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/240,960
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/239,935
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/239,937
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/241,787
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/246,474
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/246,532
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/249,216
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,210
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/226,681
 PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 60/225,759
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/225,213
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/227,182
 PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 60/225,214
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/235,836
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: 60/230,438
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/215,135
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: 60/225,266
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/249,218
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,208
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,213
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,212
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,207
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,245
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,244
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,217
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,211
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,215
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,264
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,214
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,297
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/232,400
 PRIOR FILING DATE: 2000-09-14

PRIOR APPLICATION NUMBER: 60/231,242
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,081
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,080
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,414
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,244
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/233,064
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/233,063
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,397
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,399
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,401
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/241,808
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,826
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,786
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,221
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/246,475
 PRIOR FILING DATE: 2000-11-08

Alignment Scores:

Pred. No.:	4,47e-136	Length:	951
Score:	1155.50	Matches:	226
Percent Similarity:	76.97%	Conservative:	28
Best Local Similarity:	68.48%	Mismatches:	37
Query Match:	38.09%	Indels:	41
DB:	12	Gaps:	3

US-09-601-168B-2 (1-569) x US-10-222-020-15 (1-951)

QY	15	PheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIle	34
Db	59	TTCCGAGAACACTTCAGTATGGAAGATCAAATGAAGATGAGTCCCAAG-	109
QY	35	ProGluLysAsnSerLeuArgGluThrTyrrAsnSerCysAlaArgLeuCysLeuAsnGln	54
Db	110	-----AAAATACCTTTGGCAG-	127
QY	55	GluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLys	74
Db	127	-----	127
QY	75	LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyr	94
Db	128	ATAAGTAATGGAACATCATCTGTGATTAATAATTTTGCACCATGGTCTGAATCAGATCAAGTG	187
QY	95	GluLysGluLysGluLeuCysValLysTyrPheGluGlnTrpSerGluSerAspGlnVal	114
Db	188	CAAAAAGAAAAGACTTGTGTATTAAATATTTTGCACCATGGTCTGAATCAGATCAAGTG	247
QY	115	GluPheValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSer	134
Db	248	GAATTTGTGGACATCTTATTTCACGATGTCATTATCAGCATGGACATATTAACTCT	307
QY	135	TyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAsp	154
Db	308	TACCTGAAGCCCATGTTCCAGCGGACTTTATTACCGCTTTACGAGCAGGCTTAGAT	367
QY	155	HisIleAlaGluAsnIleLeuSerTyrIleuAspAlaLysSerLeuCysAlaAlaGluLeu	174
Db	368	CACATAGCAGAAAACATCTTTCTGCTGATCCAGGTCTCTGTGTGCAGCAGAGCTG	427

175 ValCysLysGluTyrArgValThrSerAspGlyMetLeuTyrLysLeuGlu 194
 b 428 GTATGTAAGATGCGGAGTGTCTCAGAGGAATGCTTTGGAGAGCTGATTGAA 487
 y 195 ArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGln 214
 b 488 CGAATGGTACGCACTGATCCCTATGGAAGAGCTTTTCAGAAAGAGAGGTGGATCAG 547
 y 215 TyrLeuPheLysAsnLysProAspGlyAsnAlaProProAsnSerPheTyrArgAla 234
 b 548 TACCTGTTTAAAGACAGACCCACAGATGC-----CCTCCAAATTCATTTATAGTCA 601
 y 235 LeuTyrProLysIleGlnAspIleGluThrLysLeuGluThrLysLeuTyrArgVal 254
 b 602 TTATACCCAAAGATTATCCAGGATATAGAGACTATAGAAATCTAACTGGCGGTGGACGA 661
 y 255 HisSerLeuGluArgGlnHisCysArgSerGluThrLysGlyValTyrCysLeuGln 274
 b 662 CAACTTCAGAGGATTCAGTCCGCTCTGAAATAGTAAGGTGCTACTGTTTACAG 721
 y 275 TyrAspGlnLysIleValSerGlyLeuArgAspAsnThrLysIleTyrAspLys 294
 b 722 TAGCATGATGAAAAAATATCAGTCCGCTCTGAAATAGTAAGGTGCTACTGTTTACAG 780
 y 295 AsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGln 314
 b 781 ACCWGCCTGGAATGTTTGAAGTGTWAGAGGTCTTACAGGACACACAGCTCTGK-CTCTGCTGCAG 839
 y 315 TyrAspGluArgValIleLeuThrGlySerAspSerThrValArgValTyrAspVal 334
 b 840 TATGATGAKCCGTGTTTACTTACTGCTTACAGATTCACGGNGAGAGTGGGATGTG 899
 y 335 AsnThrGlyGluMetLeuAsnThrLeuLeu 344
 b 900 ACACCGGTGGAAGTCTTAACACATTTGATC 929
 RESULT 8
 S-10-116-016-15
 Sequence 15, Application US/10116016
 Publication No. US20030054379A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PT208C1
 CURRENT APPLICATION NUMBER: US/10/116,016
 PRIORITY FILING DATE: 2002-04-05
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 53
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 15
 LENGTH: 951
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (883)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (913)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (931)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (935)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (941)
 OTHER INFORMATION: n equals a,t,g, or c
 S-10-116-016-15
 Alignment Scores:
 red. No.: 4.47e-136 Length: 951

Score: 115.50 Matches: 226
 Percent Similarity: 76.97% Conservative: 28
 Best Local Similarity: 68.48% Mismatches: 37
 Query Match: 38.09% Indels: 41
 DB: 14 Gaps: 3
 US-09-601-168B-2 (1-569) x US-10-116-016-15 (1-951)
 QY 15 PheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIleIle 34
 Db 59 TTCAGAACACTTCAGTTATGGAAGATCAAAATGAAGATGAGTCCCAAG----- 109
 QY 35 ProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGln 54
 Db 110 -----AAAATACTCTTTGGCAG----- 127
 QY 55 GluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLys 74
 Db 127 ----- 127
 QY 75 LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyr 94
 Db 128 ATAAAGTAATGGAACATCATCTGTGATCGTCTCCAGAAAGAGGCCATCAGAGGAACTAT 187
 QY 95 GluLysGluLysGluLeuCysValLysTyrPheGluGlnTyrSerGlnSerAspGlnVal 114
 Db 188 CAAAAGAAAAAGAACTTGTGTATTAATATTTGACCAAGTGTCTGAATCAGATCAAGTG 247
 QY 115 GluPheValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSer 134
 Db 248 GAATTTGTGGACATCTTATTTACGNAATGTGTTCATATCAGCATGGCATATTAACTCT 307
 QY 135 TyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgLysLeuAsp 154
 Db 308 TACCTGAAGCCCATGTTGCGAGCGGACTTTTATTAACCGCTTTACAGAGCAAGCTTAGAT 367
 QY 155 HisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeu 174
 Db 368 CACATAGCAGAAACATCTTTTGTACCTGGATGCCAGGTCTCTGTGTGCAGCAGAGCTG 427
 QY 175 ValCysLysGluTyrArgValThrSerAspGlyMetLeuTyrLysLeuIleGlu 194
 Db 428 GTATGTAAGAAATGGCAGCAGTGTCTCAGAAAGAAATGCTTTGGAGAGCTGATTGAA 487
 QY 195 ArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGln 214
 Db 488 CGAATGGTACGCACTGATCCCTATGGAAGAGCTTTTCAGAAAGAGAGGTGGATCAG 547
 QY 215 TyrLeuPheLysAsnLysProAspGlyAsnAlaProProAsnSerPheTyrArgAla 234
 Db 548 TACCTGTTTAAAGACAGACCCACAGATGCG-----CCTCCAAATTCATTTATAGTCA 601
 QY 235 LeuTyrProLysIleLeuGluAsnIleGluThrLysLeuSerAsnThrArgCysGlyArg 254
 Db 602 TTATACCCAAAGATTATCCAGGATATAGAGACTATAGAAATCTAACTGGCGGTGGACGA 661
 QY 255 HisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGln 274
 Db 662 CACAACTTCAGAGGATTCAGTCCGCTCTGAAATAGTAAGGTGCTACTGTTTACAG 721
 QY 275 TyrAspGlnLysIleValSerGlyLeuArgAspAsnThrLysIleLysIleThrAspLys 294
 Db 722 TACGATGATGAAAAAATATCAGTCCGCTCTGAAATAGTAAGGTGCTACTGTTTACAG 780
 QY 295 AsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGln 314
 Db 781 ACCWGCCTGGAATGTTTGAAGTGTWAGAGGTCTTACAGGACACACAGCTCTGK-CTCTGCTGCAG 839
 QY 315 TyrAspGluArgValIleLeuThrGlySerAspSerThrValArgValTyrAspVal 334
 Db 840 TATGATGAKCCGTGTTTACTTACTGCTTACAGATTCACGGNGAGAGTGGGATGTG 899
 QY 335 AsnThrGlyGluMetLeuAsnThrLeuLeu 344

b 900 ACACCGGTGAAGACGCTTAACACATTGATC 929

RESULT 9

S-09-864-761-2778/c

Sequence 2778, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Shazron G.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2778

LENGTH: 479

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AP000252.1

OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 3.4

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7

Alignment Scores:

Pred. No.: 2.13e-77 Length: 479
Score: 689.00 Matches: 140
Percent Similarity: 90.12% Conservative: 6
Best Local Similarity: 86.42% Mismatches: 13
Query Match: 22.71% Indels: 3
DB: 9 Gaps: 1

US-09-601-168B-2 (1-569) x US-09-864-761-2778 (1-479)

QY 360 MetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerP-othrAsp 379
|||
Db 478 ATGGTGAACCTGTTCCAGGACCGCTCCATTGCTGTGTGGGACATGCTCTCGGACCGAT 419
|||
QY 380 IleThrLeuArgValLeuValGlyHisArgAlaValValAsnValValAspPheAsp 399
|||
Db 418 ATCACTTTTACGTGGTGTCTGCTGCTCCACTGAGCTGCTCAACGTAGTAGACTTTGAC 359
|||
QY 400 AspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThr 419
|||
Db 358 GAGAAATACATCGTGTCTGCTCTGGTGACAGGACCATCAAGCTCTGGAGCATGACACC 299
|||
QY 420 CysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAla-CysLeuGlnTyrAr 439
|||
Db 298 TGTGAATTGCTGCTACTCTCAATGGGCAACGCA-----GTCTTGTCTCCAGTACAG 245
|||
QY 439 GAspArgLeuValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCy 459
|||
Db 244 GATGGGCTGGTGTAGTGGATCATCAGATCATACCATAGGCTCTTGATATTGAATG 185
|||
QY 459 sGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAs 479
|||
Db 184 TGGCTCTCTGTTTAAAGAGTCTTAGAGGACATGAAGATTGTCGCGATGCATCCAGTTTGA 125
|||
QY 479 pAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAl 499
|||
Db 124 TAACAAGAGGATGTCAGCCGGGCTATGATGGAAATTAAGTTTAGGACTTGCACG 65
|||
QY 499 aAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSe 519
|||
Db 64 CGCTCTTGACCTCGAGCCGCCAGCAGCATTGTGTCGACATTGTGTCGACATTGTCGACATT 5
|||
QY 519 rGly 520
|||
Db 4 TGA 1

RESULT 10

US-09-213-888-20

Sequence 20, Application US/09213888A

Patent No. US20020164683A1

GENERAL INFORMATION:

APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe

APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company

TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

TITLE OF INVENTION: Encode Them

FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/213,888A

CURRENT FILING DATE: 1998-12-17

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 20

LENGTH: 1881

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: 6 myc tagged

OTHER INFORMATION: homo sapiens

US-09-213-888-20

Alignment Scores:

Pred. No.: 9.05e-71 Length: 1881

US-09-864-761-2778

OTHER INFORMATION: homo sapiens
US-09-328-877A-20

Alignment Scores:			
Pred. No.:	9,05e-71	Length:	1881
Score:	644.50	Matches:	169
Percent Similarity:	48.48%	Conservative:	103
Best Local Similarity:	30.12%	Mismatches:	222
Query Match:	21.24%	Indels:	67
DB:	10	Gaps:	15

US-09-601-168B-2 (1-569) x US-09-328-877A-20 (1-1881)

QY	8	LeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArgGluAspCysAsnAsnGly	27
DB	157	ATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGAATGAATGGAGACTTGGGGACCTC	216
QY	28	GluProProArgLysIleProGluLys-----AsnSerLeuArgGln-----Thr	43
DB	217	ACCATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGAATTCATGAAGAAAGAGTTGGAC	276
QY	44	TyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAla	63
DB	277	CAIGTTCTGAGCTCGCTCTTTCTTGGAAAGAAACCATCAAGTCTCAGAAATAT	336
QY	64	MetLysThrGluAsnCysVal-----AlaLysThrLys	74
DB	337	ACAAGTACCCTCGGCTTGTACCATGTTTCAGCAACACCAACCACTTTTGGGGACCTCAGA	396
QY	75	LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyr	94
DB	397	GCAGCCAAATGGCCCAAGG-----CAACAACGACCGCAATATACATCTGTC	441
QY	95	GluLysGluLysGluLeu-----CysValLysTyrPheGlnThrTrpSerGluSerAsp	112
DB	442	CAGCCACCTACAGGCTCCAGGAATGGCTAAATGTTTCAGAGCTGGAGTGGACAGAG	501
QY	113	GlnValGluPheValGluHisLeuSerGlnMetCysHisTyrGlnHisGlyHisIle	132
DB	502	AAATGCTTGTCTTAGATGAACATTCATATGATGTTGTGAACCAACACCAAGTAAACATATG	561
QY	133	AsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGly	152
DB	562	ATGCAAGTGTATAGAACCCAGTTTCAACGAGACTTCATTTCTGCTCCCT	612
QY	153	LeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAla	172
DB	613	---AAAGAGTTGGCACTCTATGTGCTTTTCATTCCTGGAACCAACCAAGACCTGCTACAAGCA	669
QY	173	GluLeuValCysLysGluTyrArgValThrSerAspGlyMetLeuTyrLysLeu	192
DB	670	GTCAGACATGCTGCTACTGAGAAATTTGGCTGAAGACCAACCTTCTCTGGAGAGAGAA	729
QY	193	IleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGlu-----	208
DB	730	TGCAAGAA-----GAGGGGATTGATGAACCATTCGCATC	765
QY	209	---ArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAlaPro	227
DB	766	AAGAGAAGAA-----GTAATAAACCCAGGTTTCATACACAGTCCA	807
QY	228	ProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGlu	247
DB	808	TGAAAGAGTGCATAC-----ATCAGACAG-----CACAGAATTGAT	843
QY	248	SerAsnTyrArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSer	267
DB	844	ACTAACTGAGCGGAGGAGAACTCAATCTCCTAAGGTG---CTGAAAGACATGATGAT	900
QY	268	LysGlyValTyrCysLeuGlnTyrAspGlnLysIleValSerGlyLeuArgAspAsn	287
DB	901	CATGTGATCACAATGTTTACAGTTTGTGGTAAACCAATGATGTTGTTCTGATGACAAC	960

RESULT 12

US-10-245-618-13
; Sequence 13, Application US/10245618
; Publication No. US20030143582A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; APPLICANT: Strohmaier, Helmo
; APPLICANT: Spruck, Charles
; APPLICANT: Sangfelt, Olle
; TITLE OF INVENTION: HCDC4 MODULATES CYCLIN E DEGRADATION
; FILE REFERENCE: TSRI 779.2
; CURRENT APPLICATION NUMBER: US/10/245,618
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/404,116
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/322,947
; PRIOR FILING DATE: 2001-09-14

QY	288	ThrIleLysIleTyrAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThr	307
DB	961	ACTTTAAAGATTTGGTCAGCAGTCAAGCAAAATGTCGAGAACATTTAGTGGACATACA	1020
QY	308	GlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSer	327
DB	1021	GGTGAGTATGGTCATCACAAATCAGGACAAACATCATCTATTAGTGGATCTACAGATCGG	1080
QY	328	ThrValArgValTyrAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCys	347
DB	1081	ACACTCAAGTGTGGAATTCAGAGACTGGAGAATGTATACACACCTTATATGGCATACT	1140
QY	348	GluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArg	367
DB	1141	TCCACTGTCGCTTGTATGTCATCTTCAAAAAAGAGTGTGTAGCGGTCTTCAGATGCC	1200
QY	368	SerIleAlaValTyrAspMetAlaSerProThrAspIleThrLeuArgValLeuVal	387
DB	1201	ACTCTTAGGGTTCGGATATTGACACAGGCCAGTGTTA-----CATGTTTGTATG	1251
QY	388	GlyHisArgAlaAlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSer	407
DB	1252	GGTCATGTTGCGAGCTCCGCTGTGTTCAATATGATGGCAGGAGGTGTGTAGTGAGCA	1311
QY	408	GlyAspArgThrIleLysValTyrAsnThrSerThrCysGluPheValArgThrLeuAsn	427
DB	1312	TATGATTTTATGTTAAAGTGTGGATCCAGAGACTGAAACCTGTCTACACACGTTGCAG	1371
QY	428	GlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValSerGlySer	447
DB	1372	GGGCATATACTAATAGAGTCTATTTCATACAGTTTGTATGATGATGATGATGATGATCT	1431
QY	448	SerAspAsnThrIleArgLeuTyrAspIleGluCysGlyAlaCysLeuArgValLeuGlu	467
DB	1432	CTTGATACATCCATCCGCTGTTTGGATGTGGACAGAGGATTCGATTCACACGTTAACA	1491
QY	468	GlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyAla	487
DB	1492	GGGACACCACTGTTTAAACAGTGAATGGAATGGAATCAAAACAAATATCTTCTCTGGGAAT	1551
QY	488	TyrAspGlyLysIleLysValTyrAspLeuValAlaAlaLeuAspProArgAlaProAla	507
DB	1552	GCAGATTCACAGTTAAATCTGGATATCAAAACAGGA-----	1590
QY	508	GlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPheArg	524
DB	1591	-----CAGTGTTTTACAAACATTTGCAAGGTCCTCAAGCATCAGAGTGTGTGACCTGT	1644
QY	525	LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTyr	544
DB	1645	TTCAGTTCAACAGAACTTTGTATTTACCGCTCAGATGTGGAACCTGTAAACCTATGG	1704
QY	545	Asp	545
DB	1705	GAC	1707

NUMBER OF SEQ ID NOS: 48

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13

LENGTH: 1620

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthesized

US-10-245-618-13

Alignment Scores:

red. No.: 2.63e-70 Length: 1620
Score: 640.00 Matches: 167
Percent Similarity: 47.79% Conservative: 92
Best Local Similarity: 30.81% Mismatches: 211
Query Match: 21.09% Indels: 72
Gaps: 15

S-09-601-168b-2 (1-569) x US-10-245-618-13 (1-1620)

14 LysPheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIle 33
10 AAGTTGGACCATGGTCTCGAGGTGGCTCTTTCTTTGGGAAGAAACCATGCAGATC 69
34 IleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsn 53
70 -----TCGAATATACAAGTACCCTGGGCTT----- 96
54 GlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThr 73
97 -----GTACCATGTTACGCA-----ACACCAACAACCTTTGGGACCTC 135
74 LysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSer 93
136 AGAGCAGCCCAATGGCCAAAGG-----CAACACGACGCCGAATTACATCT 180
94 TyrGluLysGluLysGluLeu-----CysValLysTyrPheGlnGlnTrpSerGluSer 111
181 GTCCAGCCACCTACAGGCTCCAGGAATGGCTAAATGTTTCAGAGCTGAGTGGACCA 240
112 AspGlnValGluPheValGluHisLeuSerGlnMetCysHisTyrGlnHisGlyHis 131
241 GAGAAATTGCTTGTCTTAGATGAACCTCATTCATAGTTGTGAACCAACACCAAGTAAACAT 300
132 IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg 151
301 ATGATGCAAGTCATAGAACCCAGTTTCAACGAGACTTCATTTCATGCTCCCT----- 354
152 GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla 171
355 -----AAAGAGTTGGCACTCTATGTGCTTTTCATTCCTGGAAACCAAGACCTGCTACAA 408
172 AlaGluLeuValCysLysGluTyrTyrArgValThrSerAspGlyMetLeuTyrLysLys 191
409 GCAGCTCAGACATGCTGCTACTGGAATTTGGCTGAAGCAACCTTCTCTGGAGAG 468
192 LeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGlu----- 208
469 AAATGCAAGAA-----GAGGGGATTGATGAACCATTCGCAC 504
209 -----ArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAla 226
505 ATCAGACAGAA-----GTAATAAACCCAGGTTTCATACACAGT 546
227 ProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIle 246
547 CCATGGAAGTGCATAC-----ATCAGACAG-----CACAGATT 582
247 GluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThr 266
583 GATACATACTGAGCGGAGGAACCTCAAAATCTCCTAAGGTG---CTGAAAGACATGAT 639
267 SerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAsp 286

Db 640 GATCATGTGATCATGCTTACAGTTTCTGTGTACCGAATAGTTAGTGGTTCTCATGAC 699
Qy 287 AsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHis 306
Db 700 AACACATTAAAGTTTGGTCAGCAGTCAACAGCAAAATGTCTGAGAACATTAGTGGACAT 759
Qy 307 ThrCysSerValLeuCysLeuGlnTyrAspGluArgValIleIleThrCysSerSerAsp 326
Db 760 ACAGTGGAGTATGGTCATCACAATGAGAGACACATCATCAITAGTGATCTACAGAT 819
Qy 327 SerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHis 346
Db 820 CGGACACTCAAGTGTGAATGCAGAGACTGGAGAATGTATACACACCTTATATGGGAT 879
Qy 347 CysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAsp 366
Db 880 ACTTCCACTGTGCGTTGTATGCATCTTCATGAAAAAGAGTGTGTAGCGGTTCTCGAGAT 939
Qy 367 ArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeu 386
Db 940 GCCACTCTTAGGGTTGGGATATTGAGACAGAGCCAGTGTTA-----CATGTTTG 990
Qy 387 ValGlyHisArgAlaValAlaValAsnValValAspPheAspAspLysTyrIleValSerAla 406
Db 991 ATGGGTCAATGTTCCAGCAGTCCGCTGTGTTCAATATGATGGCAGGAGGTTGTAGTGA 1050
Qy 407 SerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeu 426
Db 1051 GCATATGATTTTATGATAAGGTGTGGATCCAGAGACTGAAACCTGTCTACACAGTTG 1110
Qy 427 AsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValSerGly 446
Db 1111 CAGGGGCACTACTAATAGACTCTATTCAATACAGTTTGTATGGTATCCATGTGGTGTAGTGA 1170
Qy 447 SerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyValaCysLeuArgValLeu 466
Db 1171 TCCTTGTATACATCAATCCGTGTTGGATGGAGACAGGGAATTCATTACACAGTTA 1230
Qy 467 GluGlyHisGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGly 486
Db 1231 ACAGGGCAGCAGTCGTAAACAGTGAATGGAACCTCAAGACATAATTCCTGTCTCTGG 1290
Qy 487 AlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaPro 506
Db 1291 AATGCAGATTCTACAGTTAAATCTGGGATATCAAAACAGGA----- 1332
Qy 507 AlaGlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPhe 523
Db 1333 -----CAGTGTTTACAAACATTGCAAGTCCCAAGCATCAGAGTGTCTGTGACC 1383
Qy 524 ArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIle 543
Db 1384 TGTTACAGTTCAACAGAACTTGTATTACCACTCAGATGATGGAACCTGTAAACTA 1443
Qy 544 TrpAsp 545
Db 1444 TGGGAC 1449

RESULT 13

US-10-245-618-7

; Sequence 7, Application US/10245618

; Publication No. US20030143582A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven

; APPLICANT: Strohmater, Heimo

; APPLICANT: Spruck, Charles

; APPLICANT: Sangfelt, Olle

; TITLE OF INVENTION: HCD4 MODULATES CYCLIN E DEGRADATION

; FILE REFERENCE: TSRI 779.2

; CURRENT APPLICATION NUMBER: US/10/245,618

; CURRENT FILING DATE: 2002-09-16

; PRIOR APPLICATION NUMBER: US 60/404,116

CURRENT APPLICATION NUMBER: US/10/245,618
 CURRENT FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US 60/404,116
 PRIOR FILING DATE: 2002-08-15
 PRIOR APPLICATION NUMBER: US 60/322,947
 PRIOR FILING DATE: 2001-09-14
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 17
 LENGTH: 1770
 TYPE: DNA
 ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(1770)
 S-10-245-618-17

Alignment Scores:
 Seq. No.: 3,04e-70 Length: 1770
 Core: 640.00 Matches: 167
 Percent Similarity: 47.79% Conservative: 92
 Best Local Similarity: 30.81% Mismatches: 211
 Very Match: 21.09% Indels: 72
 Gaps: 15

S-09-601-168b-2 (1-569) x US-10-245-618-17 (1-1770)

14 LysPheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProArgLysIle 33
 157 AAGTTGGACCATGTTCTGAGTCCGCTCTTTCTTTGGGAAAGAACCATGCAAGTC 216
 34 IleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeu 53
 217 -----TCGATATACAGTACCACTCGGCTT----- 243
 54 GlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThr 73
 244 -----GTACCATGTTTCAGCA-----ACACCAACAATTTTGGGACCTC 282
 74 LysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSer 93
 283 AGACGACCAATGGCCCAAGG-----CAACACGACGCCGAATTCATCT 327
 94 TyrGluLysGluLysGluLeu-----CysValLysTyrPheGluGlnTrpSerGluSer 111
 328 GTCCAGCCACCTACAGGCTCCAGGATGGCTAAATGTTTCAGAGCTGGAGTGACCA 387
 112 AspGlnValGluPheValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHis 131
 388 GAGAAATGCTTGTCTTAGATGAACCTCATTTGATGTTGTAACCAACAGTAAACAT 447
 132 IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg 151
 448 ATGATGCAAGTGATAGAACCCACCTTTCACAGAGACTTCATTTCTGCTCCCT----- 501
 152 GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla 171
 502 -----AAAGAGTTGGCACTCTATGTGCTTTTCATCTCTGGAACCAAGACCTGTACAA 555
 172 AlaGluLeuValCysLysGluTyrTyrArgValThrSerAspGlyMetLeuTrpLysLys 191
 556 CGAGCTCAGACATGCTGCTAGAGAAATTTGGCTGAAGCAACCTTCTCTGGAGAGAG 615
 192 LeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGlu----- 208
 616 AAATCAAAAGAA-----GAGGGGATTCATGAACCATTCGCAC 651
 209 -----ArgArgGlyTyrTrpGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAla 226
 652 ATCAAGAGAGAGAAA-----GTAAATAAACCCAGGTTTCATACACAGT 693
 227 ProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIle 246

Db 694 CCATGGAAAGTGCATAC-----ATCAGACAG-----CACAGATT 729
 Qy 247 GluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThr 266
 Db 730 GATACTAACTGGAGCGAGAGAACTCAATCTCTAAGGTG--CTGAAGAGCATGAT 786
 Qy 267 SerLysGlyValTyrCysLeuGlnTyrAspGlnLysIleValSerGlyLeuArgAsp 286
 Db 787 GATCATGTGATCACATGCTTACAGTTTGTGTGAACCGAATAGTTAGTGTCTGTGATGAC 845
 Qy 287 AsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHis 306
 Db 847 AACACTTTAAAGTTTGGTCACAGCTCAGGCAAAATCTCTGAGAACAATTAGTGGGCAT 906
 Qy 307 ThrGlySerValLeuCysLeuGlnTyrAspGlnValIleIleThrGlySerSerAsp 326
 Db 907 ACAGTGGAGTATGTCATCACAAATGAGACACACATCATCATTAGTGATCTACAGAT 966
 Qy 327 SerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHis 346
 Db 967 CGGACACTCAAGTGTGGAATCGAGAGACTGGAGAATGTATACACACCTTATATGGGCAT 1026
 Qy 347 CysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAsp 366
 Db 1027 ACTTCCACTGTGGTGTGATGATCTTCTGAAAAGAGTTGTAGCGGTCTTCGAGAT 1086
 Qy 367 ArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeu 386
 Db 1087 GCCACTCTTAGGTTTGGGATATTGAGACAGGCCAGCTGTTTA-----CATGTTTGT 1137
 Qy 387 ValGlyHisArgAlaAlaValAsnValValAspPheAspAspLysTyrIleValSerAla 406
 Db 1138 ATGGGTCATGTTGCAGCAGTCCGCTGTTTCATATGATGCAGAGGGTGTGTAGTGA 1197
 Qy 407 SerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeu 426
 Db 1198 GCATGATTTTATGTAAGGTGTGGATCCAGAGACTGAAACCTGCTACACACGTTG 1257
 Qy 427 AsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGly 446
 Db 1258 CAGGGGCATCTAATAGACTCTATTCTATACAGTTTGTATGATGATCCATGCTGTGAGTGA 1317
 Qy 447 SerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyValaCysLeuArgValLeu 466
 Db 1318 TCCTTTGATACATCAATCGTGTGTGGGATGGGAGACAGGGAATTGCATTCACACGTTA 1377
 Qy 467 GluGlyHisGluGluLeuValArgCysLysLeuArgPheAspAsnLysArgIleValSerGly 486
 Db 1378 ACAGGCAACCCAGTCGTTAACAAGTGAATGGAACTCAAGACAATAATTTCTGTCTCTGGG 1437
 Qy 487 AlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaPro 506
 Db 1438 AATGCAGATCTACAGTTAAATCTGGGATATCAAAACAGGA----- 1479
 Qy 507 AlaGlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPhe 523
 Db 1480 -----CAGTGTTTACAAACATTGCAAGGTCCCAACAAGCATCAGAGTGTCTGTGACC 1530
 Qy 524 ArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIle 543
 Db 1531 TGTTTACAGTTCAACAGAACTTTGTAATTACCAGCTCAGATGATGGAACTGTAAACTA 1590
 Qy 544 TrpAsp 545
 Db 1591 TGGGAC 1596

RESULT 15

US-10-245-618-5
 ; Sequence 5, Application US/10245618
 ; Publication No. US20030143582A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven
 ; APPLICANT: Strohmaier, Heimo

APPLICANT: Spruck, Charles

APPLICANT: Sangfelt, Olle

TITLE OF INVENTION: HCDC4 MODULATES CYCLIN E DEGRADATION

FILE REFERENCE: TSRI 779.2

CURRENT APPLICATION NUMBER: US/10/245,618

CURRENT FILING DATE: 2002-09-16

PRIOR APPLICATION NUMBER: US 60/404,116

PRIOR FILING DATE: 2002-08-15

PRIOR APPLICATION NUMBER: US

PRIOR FILING DATE: 2001-09-14

NUMBER OF SEQ ID NOS: 48

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 5

LENGTH: 1884

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: CDS

NAME/REF: CDO
LOCATION: (1)... (1884)

10-245-618-5

ianment Scores:

	Length:	1884
ed. No.:	Matches:	167
re:	Conservative:	92
Percent Similarity:	Mismatches:	211
st Local Similarity:	Indels:	72
% Local Similarity:	Gaps:	15
ery Match:		
:		

-09-601-168B-2 (1-569) x US-10-245-618-5 (1-1884)

14 LysPheMetAsnSerSerGluArgGluAspCysAsnAenGlyGluProArgGlyIle 33
::: ||||| :::: |||||
271 AAGTTGACCACATGTTTCTGAGTCCGCTCTTTTCTTTGGGAAAGAACCATCAAGTC 330
::: ||||| :::: |||||
34 IleProGluLysAsnSerLeuArgGlnThrTyraSerCysAlaArgLeuCysLeuAsn 53
||| ||||| |||||
331 -----TCAGAAATATACAGTACCCTGGGCTT----- 357
||| ||||| |||||
54 GlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThr 73
||| ||||| |||||
358 -----GTACCATGTTTCAGCA-----ACACCAACAACATTTTGGGGACCTC 396
||| ||||| |||||
74 LysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSer 93
::: ||||| |||||
397 AGAGCAGCCAAATGGCCAAAGG-----CAACAACAGCCGCAAAATTCATCT 441
||| ||||| |||||
94 TyrGluLysGluLysGluLeu-----CysValLysTyrPheGluGlnTTrpSerGluSer 111
::: ||||| |||||
442 GTCCAGCCACTACAGGCTCCAGGAATGGCTAAAAATGTTTTCAGACCTGAGATGGACCA 501
||| ||||| |||||
112 AspGlnValGluPheValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHis 131
||| ||||| |||||
502 GAGAAATGCTTCTTATAGTAACTCATGTAGTGTGTGAACCAACACAGTATAAACAT 561
||| ||||| |||||
132 IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg 151
::: ||||| |||||
562 ATGATGCAAGTATGATAGAACCCAGTTTCAACGAGACTTCATTTCATTGCTCCCT----- 615
||| ||||| |||||
152 GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla 171
||| ||||| |||||
616 -----AAAGAGTTGGCATCTATGTGCTTTCATTCCTGGAACCAAGACCTGCTACAA 669
||| ||||| |||||
172 AlaGluLeuValCysLysGluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLys 191
||| ||||| |||||
670 GCAGCTCAGACATGTCCGCTACTCGAGAAATTTTGGCTGAAGACACACTTCTCTGAGAGAG 729
||| ||||| |||||
192 LeuIleGluArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGlu----- 208
||| ||||| |||||
730 AAATGCAAGAA-----GAGGGGATTTGATGAACCATTCGCAC 765
||| ||||| |||||
209 -----ArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAla 226
||| ||||| |||||

Search completed: October 22, 2003, 18:30:55

Job time : 414 secs

.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

MM protein - nucleic search, using frame_plus_p2n model

run on: October 22, 2003, 14:52:38 ; Search time 100 seconds
(without alignments)
2511.470 Million cell updates/sec

Title: US-09-601-168B-2
Perfect score: 3034
Sequence: 1 MDPAEVLQKALPKMNSSE.....PAAQAEPSPFSRTYTVISR 569

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame_p2n.model -DEV=xlpl
Q=/cgn2_1/USPTO.spool_p/US09601168/runat_22102003_100546_17813/app.query.fasta_1.711
DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09601168@cgn2_1_1.85@runat_22102003_100546_17813 -NCPU=6 -ICPU=3
NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Match	Length	DB	ID	Description
1	520	17.1	2481	3	US-08-899-578-1	Sequence 1, Appli
2	392	12.9	2237	4	US-08-914-999-7	Sequence 7, Appli
3	365.5	12.0	2085	2	US-08-283-917-8	Sequence 8, Appli
4	365.5	12.0	2085	2	US-08-961-716-8	Sequence 8, Appli
5	337	11.1	3465	4	US-08-914-999-5	Sequence 5, Appli
6	334.5	11.0	1909	4	US-09-620-312D-730	Sequence 730, App
7	308	10.2	2359	1	US-08-188-582-4	Sequence 4, Appli
8	308	10.2	2359	1	US-08-646-715-4	Sequence 4, Appli
9	299.5	9.9	7042	3	US-09-092-508-1	Sequence 1, Appli
10	299.5	9.9	7042	4	US-09-435-115-1	Sequence 1, Appli
11	299.5	9.9	7042	4	US-09-098-310-1	Sequence 1, Appli
12	299.5	9.9	7042	4	US-09-690-364-21	Sequence 21, Appli

13	299.5	9.9	7075	3	US-09-092-508-15	Sequence 15, Appli
14	299.5	9.9	7075	4	US-09-435-115-15	Sequence 15, Appli
15	298.5	9.8	2152	1	US-08-188-582-17	Sequence 17, Appli
16	298.5	9.8	2152	1	US-08-646-715-17	Sequence 17, Appli
17	291	9.6	5152	4	US-09-690-364-10	Sequence 10, Appli
18	287	9.5	3747	4	US-09-690-364-17	Sequence 17, Appli
19	268.5	8.8	1211	3	US-09-063-743-2	Sequence 2, Appli
20	268.5	8.8	1211	4	US-09-590-540-2	Sequence 2, Appli
21	268	8.8	2272	4	US-09-108-857-1	Sequence 1, Appli
22	264.5	8.7	1738	4	US-09-620-312D-615	Sequence 615, App
23	264.5	8.7	2100	4	US-09-013-118-2	Sequence 2, Appli
24	256	8.4	3067	4	US-09-016-434-1285	Sequence 1285, Ap
25	253	8.3	7518	4	US-09-620-312D-1051	Sequence 1051, Ap
26	246.5	8.1	1115	1	US-08-190-802A-19	Sequence 19, Appli
27	246.5	8.1	1115	3	US-08-477-346-19	Sequence 19, Appli
28	246.5	8.1	1115	4	US-08-473-089-19	Sequence 19, Appli
29	246.5	8.1	1115	4	US-08-487-072A-19	Sequence 19, Appli
30	239	7.9	7881	2	US-08-751-189-1	Sequence 1, Appli
31	239	7.9	7881	2	US-08-060-836-1	Sequence 1, Appli
32	239	7.9	7881	3	US-09-184-445-1	Sequence 1, Appli
33	233	7.7	2119	3	US-09-032-372-6	Sequence 6, Appli
34	226	7.4	126	3	US-09-172-841-2	Sequence 2, Appli
35	226	7.4	126	4	US-08-951-621-2	Sequence 2, Appli
36	226	7.4	7886	2	US-08-751-189-2	Sequence 2, Appli
37	226	7.4	7886	2	US-08-060-836-2	Sequence 2, Appli
38	226	7.4	7886	3	US-08-184-445-2	Sequence 2, Appli
39	225.5	7.4	3363	4	US-09-620-312D-222	Sequence 222, App
40	222	7.3	1221	3	US-08-965-600-2	Sequence 2, Appli
41	222	7.3	1221	4	US-09-489-506-2	Sequence 2, Appli
42	220	7.3	2369	4	US-09-302-769-20	Sequence 20, Appli
43	219	7.2	1517	3	US-09-180-783-1	Sequence 1, Appli
44	217	7.2	4911	4	US-09-718-692-1	Sequence 1, Appli
45	217	7.2	4911	4	US-09-718-692-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-899-578-1
; Sequence 1, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2481 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 91..1854
 JS-08-899-578-1

Alignment Scores:

Pred. No.: 7,88e-50 Length: 2481
 Score: 520.00 Matches: 150
 Percent Similarity: 43.85% Conservative: 78
 Sest Local Similarity: 28.85% Mismatches: 222
 Query Match: 17.14% Indels: 70
 DB: 3 Gaps: 17

JS-09-601-168B-2 (1-569) x US-08-899-578-1 (1-2481)

```

2y 72 LysThrLysLeuAlaAsnGlyThrSerSerMetIle----- 83
Db 193 GAAAGCTCGPACAGCAATGCTCTTCCAGCTACAAATGCTGACAAATTATCGTCCTCC 252
2y 84 valprolysglnArglys-----LeuSerAlaSerTyrGluLysGluLysGluLeu--- 100
Db 253 AGACCTTTGCAACACAACTTGATTTATCGGCTTCTCCTCTCGAAACAACAGCCTCAAT 312
2y 101 -----CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGlu 115
Db 313 CCGCGTGTGCAACATTTGATCGCATTTATTCAGGATCTATCAAGCGCGGAACAATGGAT 372
2y 116 PheValGluHisLeuSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyr 135
Db 373 GCATTCACAGCTGTCTGAGGAATCCAAATATTCGACAGTGGGTGCCATT 432
2y 136 LeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHis 155
Db 433 ATTGAACCCCATTTCCAGCGTGATTTCTCTCTGCTCCCTGTC-----GAG 480
2y 156 IleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuVal 175
Db 481 CTGGGAATGAAATCTCTCACAAATTTAACCGGATATGACCTCTCAAAAGTGCACAGGTG 540
2y 176 CysLysGluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLeu---IleGlu 194
Db 541 TCGAAAATTTGGAAATTTGATCTGAAATTCGAAATTCGAAATTTGGAAGAGTCTTGGTGTCAA 600
2y 195 ArgMetVal-----ArgThrAspSerLeuTrpArgGlyLeuAla 207
Db 601 GAGTTTAAACATCATCCAGATCCACAGACCGAGTTTACTGGTGGCGCAAGAACTGCA 660
2y 208 GluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAlaPro 227
Db 661 ATTGCTGTGTA-----GTCACATTTCTGATCACAATTCAGCCA 699
2y 228 ProAsn---SerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIle----- 243
Db 700 TGTGATCTTAATGTTATCGATCTTAAGTTCGACAGAGTTTGGAGATATCTTGAAGCGC 759
2y 244 -----GluThrIleGluSerAsnTrpArgCysGly 253
Db 760 GCTGCTGACAAGTCACGTTATCTTCGAGCGCGATAAAATTTGAAAAGAACTGGAATCGCAAT 819
2y 254 ArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeu 273
Db 820 CCAATATGGGTGACAGTG---CTACGAGGACACGAAGATCATGTAACTACTTGTATG 876
2y 274 GlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAsp 293
Db 877 CAAATTCATGATGATGCTGGTGGTACTGATCTGACGATAACACTCTTAAAGATGTTGGT 936
2y 294 LysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeu 313
Db 937 ATTGACAAAGGAGGATTTGTATACACACTAGTCGCGCACACTGGAGGATTTGGACATCA 996

```

```

Qy 314 GlnTyrAspGlu-----ArgValIleIleThrGlySerSerAspSerThrValArgVal 331
Db 997 CAGATTCTCAATGCGGAAGATATATTGTTAGCGGTCCACTGATAGAACTGTAAAGTT 1056
Qy 332 TrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisHisCysGluAlaValLeu 351
Db 1057 TGGAGTACTGTAGATGGTTCACTTCTTCATACACTCAAGGACATACCTCCACTGTTGCA 1116
Qy 352 HisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaVal 371
Db 1117 TGCATGGCTATGCTGTTCCATATCTGCTACTGGATCACGAGATACCACTCTTCGTGTA 1176
Qy 372 TrpAspMetAlaSerProThrAspIleThrLeuArgValLeuValGlyHisArgAla 391
Db 1177 TGGAGCTAGAAATCCGAGCTCACCTGCA-----ACTTATACATGGCCATCATGCA 1227
Qy 392 AlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSerGlyAspArgThr 411
Db 1228 GCGTTTCGATGCGTTCAATTCGATGGAACAACTGTGTTTCGGGAGGATATGATTTTACC 1287
Qy 412 IleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArg 431
Db 1288 GTTAAATTTGGAATGCTCATACTCGGAGATGTATCCGCTACTCTGACCGGTCAACAAT 1347
Qy 432 GlyIleAlaCysLeuGlnTyrArgAspArg-----LeuValValSerGlySerSerAsp 449
Db 1348 AGAGTTTATCTCTTCTCTTTGAAAGCGGCGATCGTGTGCTCTGCTCTCTGGAC 1407
Qy 450 AsnThrIleArgLeuTrpAspIle-----GluCysGlyAlaCysLeuArgValLeu 466
Db 1408 ACTTCAATTCGCTGGTGGGATTTTACACGACCGGAAGCCAGAAATGTGTGGCTCTTTG 1467
Qy 467 GluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGly 486
Db 1468 CAAAGACACACTCTACATCCGGAATGCACTTCGAGGCAATATCTCGTGTGATGC 1527
Qy 487 AlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaPro 506
Db 1528 AATGCAATGAGCATGTTAGATATGGATAT-----CAC 1563
Qy 507 AlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArgLeuGln 526
Db 1564 GAGGGAAC---TGTGTACACATGCTTCTTGACATCGATCCGCTATCCTTCACTTCA 1620
Qy 527 ---PheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrpAsp 545
Db 1621 TGGTTTGGACGAAATATGTTAGTACGAGTAGTAGGAACTGTCAAATTTGGGAT 1680

```

RESULT 2

US-08-914-999-7
 ; Sequence 7, Application US/08914999
 ; Patent No. 6346406
 ; GENERAL INFORMATION:
 ; APPLICANT: RYAZANOV, Alexey G.
 ; APPLICANT: Hait, William N.
 ; APPLICANT: Pavur, Karen S.
 ; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
 ; TITLE OF INVENTION: AND METHODS OF USE THEREFOR
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

[illegible]

```

RESULT 3
US-08-283-917-8
; Sequence 8, Application US/08283917
; Patent No. 5849557
;
; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: INOUE, KEIZO
; APPLICANT: ARAI, HIROYUKI
;
; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
;
; TITLE OF INVENTION: AND GENE THEREOF
;
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
;

```

Alignment Scores:					
Pred. No.:	4.59e-35	Length:	2237		
Score:	392.00	Matches:	119		
Percent Similarity:	41.18%	Conservative:	84		
Best Local Similarity:	24.14%	Mismatches:	172		
Query Match:	12.92%	Indels:	118		
DB:	4	Gaps:	12		
JS-09-601-168B-2 (1-569) x US-08-914-999-7 (1-2237)					
Y	13	LeuLysPheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProArgLys	32		
b					
b	1048	TAAATTTCCGTGATACATCTGAAGAGATAAATAATAATAATAATAATAATAATAA	1107		
Y	33	IleIleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeu	52		
b					
b	1108	AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	1167		
Y	53	AsnGlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLys	72		
b					
b	1168	TCAAAATCAATTAGTTGAAATTTCAAAGTGCTAGTAAGAAGAAAGATGAGAGATTCGCCA	1227		
Y	73	ThrLys-----LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLys	89		
b					
b	1228	AGTAGACAATTATTGTGTTCCAAATGATGGTAATACATTAATAATACAAATAAGAGAGATCA	1287		
Y	90	LeuSerAlaSerTyrGluLysGluLys---GluLeuCysValLysTyrPheGluGlnTrp	108		
b					
b	1288	AAATCAAAATCAATAGATTTTAGAAAACCAGAAATTTTAATAAAT----	1341		
Y	109	SerGluSeraspGlnValGluPheValGluHisLeulleSerGlnMetCysHisTyrGln	128		
b					
b	1342	AAAGAGAGTAAATTTGGAAACCGATAAAA-----	1371		
Y	129	HISGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeu	148		
b					
b	1371	-----ATTGT-----	1371		
Y	149	ProLalaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSer	168		
b					
b	1372	-----TTAATGTAACACTATTAAAGGATATCATGTTTACAAAGTCAT	1410		
Y	169	LeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSerAspGlyMetLeu	188		
b					
b	1411	TTATGT-----ATTGT-----GATAATTATTA	1434		
Y	189	TrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGlu	208		
b					
b	1435	TTTTACAGATGTTCCAGATAATCCAATTAGAGTG-----	1467		

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
 ADDRESS: NEUSTADT, P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/283,917
 FILING DATE: 03-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 209943/1993
 FILING DATE: 03-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, No. 5849557man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2292-030-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2085 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: Bos taurus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 844..2073
 US-08-283-917-8

Alignment Scores:
 Pred. No.: 4,83e-32 Length: 2085
 Score: 365.50 Matches: 140
 Percent Similarity: 38.83% Conservative: 98
 Best Local Similarity: 22.84% Mismatches: 210
 Query Match: 12.05% Indels: 165
 DB: 2 Gaps: 22

US-09-601-168b-2 (1-569) x US-08-283-917-8 (1-2085)

QY 20 GluArgGluAspCysAsnAsnGlyGluProProArgLysIlelleProGluLysAsnSer 39
 DB 463 GAGCGCGCTGAC-----CCCAACCAACCATCCGAGCATCCACCCCAA 507
 QY 40 LeuArgGlnThrTyrrAsnSerCysAlaArgLeuCysLeuAsnGln-----GluThr 56
 DB 508 ATCCGGCAGGANTTTCTTTTCGCGCTCTTCAACGGAGCTGCCTTTTGAGCTT 567
 QY 57 ValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAla 76
 DB 568 GTAACACTGAGCTTCGAGGCCCTCAGCCATCTCTTCGAATCTCCCACTCGTATAGGA 627
 QY 77 AsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLys 96
 DB 628 AACGAGCTGCCTCGCTTA-----ACCTCCCAAGGTG 657
 QY 97 GluLysGluLeuCysValLysTyrrPheGluGlnTrpSerGluSerAspGlnValGluPhe 116
 DB 658 GAATGAACCTTACTTTGTAATATCTC-----684
 QY 117 ValGluHisLeuIleSerGlnMetCysHisTyrrGlnHisGlyHis-----131
 DB 1636 ACAAGGAATCAAGGCTGAGCTTCGAGAACATGAGCATGTGGTAGAATGCAATTCCTCGG 1695

DB -----CTGGTTACACGTTGGATTTCATTGTGAAAGAAATCATTTTCCCTCGGTGAA 735
 QY -----IleAsnSerTyrrLeuLysProMet 139
 DB AGCCACTTAGTGGCTTATTAAATTATAAATCCAGGGGTTGCAAGCTTTTGG-----ATT 789
 QY -----LeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsn 159
 DB TTCAGAGGAGGACATACCACTATA-----TCGAATAGCTTGTATATTACAGCAAAATG 846
 QY -----IleLeuSer-----162
 DB GTGCTGTCACAGACACAGAGATGAATAATCGAGCTATAGCAGATTATCTTCGTCA 906
 QY -----TyrrLeuAspAlaLysSerLeuCysAlaAlaGlu-----173
 DB AATGGCTACGAAGAAGCATATTAGTTTTTAAAGGAAGCTGAATTAGATATGAATGAA 966
 QY -----LeuValCysLysGluTrpTyrrArgValThrSerAsp 185
 DB GAATTAGATAGAAATATGCTGGTCTTTTGGAAAAAATGG-----ACATCTGTT 1017
 QY GlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpArgGly 205
 DB ATTAGATTACAAAAAGAGTTATGGAATTAGAAATCAAGTTAAATGAAGCAAAAGAA 1077
 QY LeuAlaGluArgArgGlyTyrrGlyGlnTyrrLeuPheLysAsnLysProProAspGly--- 224
 DB TTACGTCGGGTGACCTCTTGGTCAG-----AAAAGAGACCCAAAGAAATGGATT 1128
 QY AsnAlaProProAsnSerPheTyrrArgAlaLeuTyrrProLysIlelleGlnAspIleGlu 244
 DB CCCCCTCCACAGAAAAATATGCAATTCAGTGGTCAT-----1164
 QY ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
 DB -----AGGAGT 1170
 QY GluThrSerLysGlyValTyrrCysLeuGlnTyrrAspAspGlnLysIlelleValSerGlyLeu 284
 DB CCAGTCACTCGAGTCATTTTCCATCTCTGTGTC-----AGTGTATGTGCTCTCTGCTCA 1224
 QY ArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
 DB GAGGATGCTACAAATAAGGTGTGGGATTATGAGACTGGAGATTTTGAACTTTAAG 1284
 QY GlyHisThrGlySerValLeuCysLeuGlnTyrrAspGlu-----ArgValIleIleThr 322
 DB GGGCATACAGACTCTGTACAGGATATTTTCAGCCACACAGTGGCAAGCTTCTGGCTTCA 1344
 QY GlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThr 342
 DB TGTTCGCAGATATGACCATTAAGCTATGGGATTTTCAGGCTTTGATGATCATCAGAAC 1404
 QY LeuIleHisHisCysGluAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMet 360
 DB ATGCATGCCATGACCAATGTTTCTTCAGTACCATCATGCCCATGGAGATCATATA 1464
 QY ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 DB GTGTCCTGCTCAAGGATATAAACTATAAAATGTGGAAAGTGAACACTGCTACTGTGTG 1524
 QY ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValVal-----AspPhe 398
 DB -----AAGACATTCACAGGACACAGAGATGGGTAGCTATGTTGGTGGCGGCAAAATCAA 1575
 QY AspAspLysTyrrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSer 418
 DB GACGSCACTCTGATAGCCAGCTGTTCCAAATGACCAGACTGTGCGTGTATGGTGTAGCA 1635
 QY ThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrr 438
 DB ACAAGGAATCAAGGCTGAGCTTCGAGAACATGAGCATGTGGTAGAATGCAATTCCTCGG 1695

```
QY 439 ArgAspArg----- 441
Db 1696 GCTCCTGAAAGCTCATATCTTCCATCTCTGAAGCAACAGGATCTGAGACTAAAAAAGT 1755
QY 442 -----LeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAsp 456
Db 1756 GGCACACCTGGGCATCTTACTGTCCGATCCAGGCAAGACTATCAAGATGGGAT 1815
QY 457 IleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIle 476
Db 1816 GTCAGTACTGGCATGTGCTTATGACCTGGTGGTCAATGATACTGGGTGAGTGT 1875
QY 477 ArgPheAspAsn-----LysArgIleValSerGlyAlaTyAspGlyLysIleVal 494
Db 1876 CTGTTTCATCTGGGGGAAGTTATTTTGGTGGCTGATGACAAAGACCCCTGGCGGTG 1935
QY 495 TrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThr 514
Db 1936 TGGGATTAC-----AAGAACAGCGATGCATGAGACC 1968
QY 515 LeuValGluHisSerGlyArgValPheArgLeuGlnPheAspGluPhe-----GlnIle 532
Db 1969 CTAATCGCGATCAACACTTGTATCTCTCTGATTTCCATAAGACGCGCCCATATGTG 2028
QY 533 ValSerSerSerHisAspAspThrIleLeuIleTrpAsp 545
Db 2029 GTTACTGGCAGTGTAGATCAACAGTAAAGGTGTGGAG 2067

RESULT 4
JS-08-961-716-8
: Sequence 8, Application US/08961716
: Patent No. 5880272
: GENERAL INFORMATION:
: APPLICANT: ADACHI, HIDEKI
: APPLICANT: TSUJIMOTO, MASAFUMI
: APPLICANT: INOUE, KEIZO
: APPLICANT: ARAI, HIROYUKI
: TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
: TITLE OF INVENTION: AND GENE THEREOF
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
: ADDRESSEE: NEUSTADT, P.C.
: STREET: 1755 S. Jefferson Davis Highway, Suite 400
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,716
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/283,917
: FILING DATE: 03-AUG-1994
: APPLICATION NUMBER: JP 209943/1993
: FILING DATE: 03-AUG-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Oblon, No. 5880272man F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 2292-030-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
```

```
:
: LENGTH: 2085 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Bos taurus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 844..2073
: US-08-961-716-8
Alignment Scores:
Pred. No.: 4.83e-32 Length: 2085
Score: 365.50 Matches: 140
Percent Similarity: 38.83% Conservative: 98
Best Local Similarity: 22.84% Mismatches: 210
Query Match: 12.05% Indels: 165
DB: 2 Gaps: 22
US-09-601-168B-2 (1-569) x US-08-961-716-8 (1-2085)
QY 20 GluArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSer 39
Db 463 GAGCGCGCTGAC-----CCACCGAACCATCCGACGATCCACCCACCA 507
QY 40 LeuArgGlnThrTyAsnSerCysAlaArgLeuCysLeuAsnGln-----GluThr 56
Db 508 ATCCGCGAGGATTTCTTTCTGCGCTCGGCTCTTCAACGGGAGCTGCTTTTGACGTT 567
QY 57 ValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAla 76
Db 568 GTAACACTGAGCTTCGAGGCCCTCAGCCATTCTCTCGAATCTCCCTCCTGATAGA 627
QY 77 AsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyArgLys 96
Db 628 AACGCGAGTGCCTGCTTA-----ACCTCCGAGGTG 657
QY 97 GluLysGluLeuCysValLysTyPheGluGlnTrpSerGluSerAspGlnValGluPhe 116
Db 658 GAATGAACCTTACTTGTGAATATCTC----- 684
QY 117 ValGluHisLeuIleSerGlnMetCysHisTyGlnHisGlyHis----- 131
Db 685 -----CTGTTACAGTGGATTCATTGTGAAGAATCAITTTCCCTCGCTGAA 735
QY 132 -----IleAsnSerTyLeuLysProMet 139
Db 736 AGCCACTTAGTGGCTTATTAAATTAATAATCCAGGGGTTCGAAAGCTTTTG-----ATT 789
QY 140 LeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsn 159
Db 790 TTCAGAGAGGGGACATAACCACTATA---TCGAATAAGCTTGATATTACAGCCAAATG 846
QY 160 IleLeuSer----- 162
Db 847 GTGCTGTCCCGACAGACACAGAGTAACTAAATCGAGCTATAGCAGATTATCTCTGTTCA 906
QY 163 -----TyrlleuAspAlaLysSerLeuCysAlaAlaGlu----- 173
Db 907 AATGGCTTACGAGAGACATATTCAGTTTTTAAAGAGGAGCTGAATTAGATATGATGAA 966
QY 174 -----LeuValCysLysGluTrpTyArgValThrSerAsp 185
Db 967 GAATTAGATAAGAAATATGCTGCTCTTTTGGAAAAAATGG-----ACATCTGTT 1017
QY 186 GlyMetLeuTrpLysLeuIleGluArgMetValArgThrAspSerLeuTrpArgGly 205
Db 1018 ATTAGATTACAAAGAGGTTATGGAATTAGAATCAAAAGTTAAATGAAGCAAGAA 1077
QY 206 LeuAlaGluArgGlyTrpGlyGlnTyLeuPheLysAsnLysProProAspGly--- 224
Db 1078 TTACGTGGGTGGACCTCTTGTGTAG-----AAAGAGACCCAAAGATGGATT 1128
```

QY 225 AsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGlu 244
 Db 1129 CCGGTCACACAGAAATATGATTCAGTGGTCAT-----1164
 QY 245 ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
 Db 1165 -----AGGAGT 1170
 QY 265 GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu 284
 Db 1171 CCGAGTACTCGAGTCATTTCCATCTGCTGTC-----AGTGTAGGTCCTCTGCTTCA 1224
 QY 285 ArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
 Db 1225 GAGGATGCTACAAATTAAGGTGTGGGATATGAGACTGGAGATTTTGAACGAACCTCTTAAG 1284
 QY 305 GlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu-----ArgValIleIleThr 322
 Db 1285 GGGCATACAGACTCTGACAGGATATTTTCATTCGACCACAGTGGCAAGCTTCTGGCTTCA 1344
 QY 323 GlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThr 342
 Db 1345 TGTTCTGCAGATATGACCATTAAGCATGCGATTTTCAGGGCTTTGAATGCATCAGAAC 1404
 QY 343 LeuIleHisCysGluAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMet 360
 Db 1405 ATGATGGCCATGACCAACATGTTCTTCATGACCATCATGCCAATGGAGATCATATA 1464
 QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 Db 1465 GTGTCTCCCTCAGGGATAAACTATAAAATGTGGAGTGCAAACTGCTACTGTGTG 1524
 QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaValAsnValVal-----AspPhe 398
 Db 1525 -----AAGACATTCACAGGACACAGAGATGGTACGTATGTGGCGCCAAATCAA 1575
 QY 399 AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSer 418
 Db 1576 GAGCGCACTGTAGTACGAGCTGTTCATGACAGACTGTGGGTGTGGGTGTGATGAGCA 1635
 QY 419 ThrCysGluPheValArgThrLeuAsnGlyHisIysArgGlyIleAlaCysLeuGlnTyr 438
 Db 1636 ACAAGGAATGCAGGCTGAGCTTCGAGAACATGAGCATGTGTAGATTCATTTCTCTG 1695
 QY 439 ArgAspArg-----441
 Db 1696 GCTCCTGAAGCTCATATTTCTTCCATCTCTGAAGCAACAGGATCTGAGACTAAAAAAGT 1755
 QY 442 -----LeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAsp 456
 Db 1756 GGCAAACCTGGGCATTTCTTACTGTCCGGATCCAGGACAGACTATCAAGATGTGGAT 1815
 QY 457 IleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIle 476
 Db 1816 GTCAGTACTGGCATGTGCTTATGACCTGTGGTGCATCATCACTGGTGTGAGT 1875
 QY 477 ArgPheAspAsn-----LysArgIleValSerGlyAlaTyrAspGlyLysIleLysVal 494
 Db 1876 CTGTTCATTTCTGGGGGAAGTTATTTTGTAGTTGGCTGATGACAAAGACCTCGCGGTG 1935
 QY 495 TrpAspLeuValAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThr 514
 Db 1936 TGGGATTAAC-----AAGAACAGCGATGATGAAGACC 1968
 QY 515 LeuValGluHisSerGlyArgValPheArgLeuGlnPheAspGluPhe-----GlnIle 532
 Db 1969 CTCAATGCGATGAACACTTTGTACTCTCTGATTTCCATTAAGAGCGGCCCATATGTG 2028
 QY 533 ValSerSerHisAspAspThrIleLeuIleTrpAsp 545
 Db 2029 GTTACTGGCAGTGTAGATCAACAGATAAAGTGTGGAG 2067

RESULT 5
 US-08-914-999-5
 ; Sequence 5, Application US/08914999
 ; Patent No. 6346406
 ; GENERAL INFORMATION:
 ; APPLICANT: Ryazanov, Alexey G.
 ; APPLICANT: Hait, William N.
 ; APPLICANT: Pavur, Karen S.
 ; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
 ; TITLE OF INVENTION: AND METHODS OF USE THEREFOR
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/914,999
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 601-1-078
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3465 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Dictyostelium discoideum
 ; US-08-914-999-5
 Alignment Scores:
 Pred. No.: 2,24e-28 Length: 3465
 Score: 337.00 Matches: 79
 Percent Similarity: 50.00% Conservative: 52
 Best Local Similarity: 30.15% Mismatches: 103
 Query Match: 11.1% Indels: 28
 DB: Gaps: 7
 US-09-601-168B-2 (1-569) x US-08-914-999-5 (1-3465)
 QY 246 IleGluSerAsnTrpArgCys-----GlyArgHisSerLeuGlnArg 259
 Db 2692 GTCAATGACAAATGGAAATGTTTATACACTGTCTCAATGGTCATAGAAAATCAATGAAAGT 2751
 QY 260 IleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLys 279
 Db 2752 ATCGCTGTAAATAGT-----AATTAC 2772
 QY 280 IleValSerGlyLeuArgAspAsnThrIleLys-----IleTrpAspLysAsnThr 296
 Db 2773 ATTTCCTCATCATCCCTGATACACCATCAAGTTCATATCATTCGTAGTGAACACC 2832
 QY 297 LeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAsp 316
 Db 2833 ---AAATGATATAGACATTTGGTTGGTCACACTGGTGAAGTTAATTGTGTCGTGCCCAAT 2889

```

Qy 317 GluArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThr 336
Db 2890 GAAAAATATCTTTTCAGTGTGTAGTTATGATAAAACTATCAAGGTTTGGGATTTGCAACC 2949
Qy 337 GlyGluMetLeuAsnThrLeuIleHis---HisCysGluAlaValLeuHisLeuArgPhe 355
Db 2950 TTTAAAGAAATTAATATTGAGGGGTTCATCAAAAGTACATTAATAACATTTGGCTTTG 3009
Qy 356 AsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAla 375
Db 3010 AGTCGACGTTATCTTTTAGTGGGTGAACGATCAATCATTTACGTTTGGGATACTGAA 3069
Qy 376 SerProThrAspIleThrLeuArgArgValLeuValGlyHisArgAlaValAsnVal 395
Db 3070 ACA-----CTTAGTATGCTTTTCAATATGCAAGGTCATGAAGATTGGGTACTCTCT 3120
Qy 396 ValAspPheAspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrp 415
Db 3121 CTTCATTGTCACGCTAGTTATCTTTCTCAACCTCAAGATTAATGTCATCAAGATTGG 3180
Qy 416 AsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCys 435
Db 3181 GATCTCTCAAAATTTTCAGTTGTGTATCTCTAATCTCTAAAGGTCATTGGAATTCCTCTCAAGT 3240
Qy 436 LeuGlnTyrArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrp 455
Db 3241 TGTGTCGTAAAGATCGTTATCTATACAGTGGTTCGAAGATTAATCAATCAAGATTGG 3300
Qy 456 AspIleArgCysGlyAlaCysLeuArgValLeu---GluGlyHisGluGluLeuValArg 474
Db 3301 GATCTCGATACACTTGAATGTGTGTACACCATTCCTCAAAATCTCATTTCTTTGGGTGTAATA 3360
Qy 475 CysIleArgPheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysVal 494
Db 3361 TGTITTAATGTTTTCATAAATCAAAATCATTTTCGCTTTCGATGGTTCATTAATTAAGTT 3420
Qy 495 TrpAsp 496
Db 3421 TGGGAA 3426

```

RESULT 6

US-09-620-312D-730

Sequence 730, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyun

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: 784C1P2B

CURRENT APPLICATION NUMBER: US/09/620,312D

PRIOR FILING DATE: 2000-07-19

PRIOR FILING DATE: 2000-04-25

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

```

; SEQ ID NO 730
; LENGTH: 1909
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (147)..(1511)
US-09-620-312D-730

Alignment Scores:
Pred. No.: 1,648-28 Length: 1909
Score: 334.50 Matches: 110
Percent Similarity: 37.50% Conservative: 67
Best Local Similarity: 23.31% Mismatches: 144
Query Match: 11.03% Indels: 151
DB: 4 Gaps: 14

US-09-601-168B-2 (1-569) x US-09-620-312D-730 (1-1909)

Qy 103 LysTyrPheGluGlnTrpSerGluSerAspGlnValGluPhe----- 116
Db 156 AAGGACTTTTGAGACATGGCTTGATAACATTTCTGTACATTTCTTCTGACGCACTTG 215
Qy 117 -----ValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHis 131
Db 216 CAGAAAAAATGAACACTCTGGATCACCTGATTAGTCTGAGTGGGGCAGTCCAGCTCAGCAT 275
Qy 132 IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg 151
Db 276 CTCTCCAATAAAGCTAGAGACTCTCTCAAGGGGAGTCTCTCAAACTCCTTCCCTG--- 332
Qy 152 GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla 171
Db 333 -----GAGCTCAGTTTTTATTGTTAAATGGCTCGATCCTCAGACTTTACTCACA 383
Qy 172 AlaGluLeuValCysLysGluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLys 191
Db 384 TGCTGCTCGTCTCTAAACAGTGAATAAGTGATAAGT----- 422
Qy 192 LeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGluArgArgGly 211
Db 423 -----GCCTGTACAGAGGTGGCGCAGACTGCATGTGTAATAAATTTGGGC 464
Qy 212 TrpGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAlaProProAsnSerPhe 231
Db 465 TGG-----CAGATAGATGATTCTGTTTCAGGACGCTTTGCAC 500
Qy 232 TyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGluSerAsnTrpArg 251
Db 501 TGGAGAAGGTTTATTGGAAGGCTATTTTGAAGATGAAGCAACTGGAGGAC----- 551
Qy 252 CysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyr 271
Db 552 -----CATGAAGCCTTTGAAACCTCGTCA----- 575
Qy 272 CysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIle 291
Db 575 ----- 575
Qy 292 TrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeu 311
Db 576 -----TTAATTGGACACAGTGCAGGATGTAT 602
Qy 312 CysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThrValArgVal 331
Db 603 GCACCTTTACTACAAAGATGGACTTCTCTGACAGGTCAGATGACTTGTCTGCAAGCTG 662
Qy 332 TrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaValLeu 351
Db 663 TGGGATGTGACGACGAGGCGAGTGGCTTTATGGCATCCAGACCCAC----- 707
Qy 352 HisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaVal 371
Db -----

```

708	----	-ACTGTGCA-----	716
372	TpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGlyHisArgAla	391	
716	----	-----	716
392	AlaValIleValValAspPheAspAspLysTyrIleValSerAlaSerGlyAspArgThr	411	
717	-----GCGGTGAAGTTTGATGAACAGAGAAGCTTGTGACAGGCTCTCTTGACAACACT	767	
412	IleLysValTtpAsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArg	431	
768	GTGCTGCTGGTAATGAGATTGGAGCCAGGACCCAGCACTTTCGGGGGCACACGGGG	827	
432	GlyIleAlaCysLeuGlnTyrArgAspArgLeu-----ValValSerGlySerSerAsp	449	
828	CGCGTATTAGCTGGGACTACAATGATGAATCTGGTGGTGGCGGCTCTGCAGAC	887	
450	AsnThrIleArgLeuTtpAspIleGluCysGlyAlaCysLeuArgValLeuGluGluHis	469	
888	TTCACCTGTGAAGATATGGCTTTATCTCTGGGACATGCGCTGAACACACTCACCGGGCAC	947	
470	GluGlu-----LeuValArgCys-----IleArgPhe	478	
948	ACGGAATGGGTACCAAGTAGTATTGCAGAGTGCAAAGTCAAGTCTCTCTGCACAGT	1007	
479	AspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTtpAspLeuVal	498	
1008	CCTGGAGACTACATCCTCTTAAGTGCAGACAAATATGAGATTAAGATTGGCCAATTGGG	1067	
499	AlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHis	518	
1068	AGAGAAATCAACTGTAAAG-----TGCTTTAAAGACATTGCTGTGCTCTCT	1109	
519	SerGlyArgValPhe-----ArgLeuGlnPheAspGluPheGlnIleValSer	534	
1110	GAGGATAGAGTATCTGCCTGCAGCCAGACTTCATTTTGATGGCAAAATACATTGCTGTG	1169	
535	SerSerHisAspAspThrIleLeuIleTtpAspPhe	546	
1170	AGTTTCAGCACTTGGT---CTTACAGTGGGACTTT	1202	

RESULT 7

US-08-188-582-4
; Sequence 4, Application US/08188582
; Patent No. 553410
; GENERAL INFORMATION:
; APPLICANT: Tian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Rupert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188.582
; FILING DATE: 28-JAN-1994

Percent Similarity: 33.77% Conservative: 102
 Best Local Similarity: 20.32% Mismatches: 195
 Query Match: 9.87% Indels: 308
 DB: 3 Gaps: 34

US-09-601-168B-2 (1-569) x US-09-092-508-1 (1-7042)

```

QY 2 AspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArg 21
DB 1769 GAAGAAGTTGAAGACATACTGCAGAG-----TTGTAAATAAGTCTCTTTA 1816
QY 22 GluAspCysAsnGlyGluProArg-----LysIle 33
DB 1817 TTCTGTGATCGGAATGGAAGTCGTTTGTATTATTATACATGATCTTCAAGTAGATT 1876
QY 34 IleProGluLysAsnSerLeuArg----- 41
DB 1877 CTTACAGAGAAGAAATGACGAGCGCTTCAGGATCTACATAAGAGATAATCACTCACTTT 1936
QY 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAla 61
DB 1937 CAGAGATATCACCAGCGCATACTCTTTCACCAGATCAGGAGACTGTATGTATGTATC 1996
QY 61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 81
DB 1997 AACTTT-CTGGCCTATCAC----- 2014
QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
DB 2015 -----ATGCCAGTCCAGAGATGACAGAGACTTTG 2046
QY 101 sValLysTyrPheGlu-----GlnTyrSerGluSerAspGlnValGluPheVal 118
DB 2047 TGCCTTAATGTTTCCCTCGATGTAATTAAGCA---AAAACAGAACTTGTAGCCCTGC 2103
QY 118 uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLys 138
DB 2104 TCATCTGATTCATGAATTTGGATACAGACAT-----AT 2139
QY 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAla 158
DB 2140 ACTAGATGAAAGGATTTGTGCA-----GTCAGTGA 2169
QY 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
DB 2170 GAAATTTTCAGAGTTTATCTTTAAATGGACACTTCTGGACGACAGCCATTCCTAA 2229
QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr 180
DB 2230 TATTGTACAACTGGGTCTCTGTGAGCCGGA-----ACTTCAGAAAGTTTATCAGCAAGC 2283
QY 181 -----ArgValThrSerAspGlyMetLeuTyr-----Ly 190
DB 2284 TAAGCTGCAGGCCAAGCAGGAGGTGTAATGGAATGCTTACCTGGAATGGATAAACAA 2343
QY 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTyrPar 204
DB 2344 AAAAACAACATCAGCAATCTTTCCCGCTTGTGTCGCCGCCACACAGATGCTGT 2398
QY 204 gGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProAspG 224
DB 2398 ----- 2398
QY 224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIle 244
DB 2399 -----TACCATGCTGCTTTCT-----CAGGATGTC 2427
QY 244 urhIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIleHisCysArg 264
DB 2428 GAGAATAGTCTT-----TGAGAGCTGATAAAACCTTACAGGTGTC---AAAGC 2475
QY 264 rGluThrSerLys-----GlyValTyrCysLeuG 274
  
```

```

DB 2476 TGAACAGGAGAGAAACTTCTAGAAATCAAGGCTCATGAGGATGAAGTGTCTTGTGTC 2535
QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTy 292
DB 2536 ATTCTCTACAGATGACAGATTATTACCAACTGCTCAGTGGATGATAAAAAAGTGAAGATTG 2595
QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
DB 2596 GAATTCCTATGACTGGGAACTAGTACACACCTATGATGAGCACTCAGAGCAAGTCAATTG 2655
QY 312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh 328
DB 2656 CTGCCATTTCCACCAACAGTAGTCATCTCTCTAGCCACTGGTCAAGTGCCTT 2715
QY 328 rValArgValTyrAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysG 348
DB 2716 CCTCAAACTTTGGATTTGAATCAAAAAGAATGTGCAAAATACCATGTTTGTGCATCAA 2775
QY 348 uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366
DB 2776 TTCAGTCAATCACTGCAGATTTCACAGATGATGAAGCTTTTGGCTAGTGTTCAGCTGA 2835
QY 366 pArgSerIleAlaValTyrAspMetAlaSerProThrAsp-----IleThrLeuAr 383
DB 2836 TGAACCTTAAAGCTTTGGGATGCGACATCAGCAAAATGAGAGGAAAAGCATTAATGTGA 2895
QY 383 g----- 383
DB 2896 ACAGTCTTCTTAATTTGGAGGACCTCAAGAGATATGGAAGTAGTAGTGAAGTGTG 2955
QY 384 -----ArgValLeuVal 387
DB 2956 TTCGTGCTGCTGATGTTGAGGATGATATGTTGTCGACGACAAAATAAAATCTTTTGTG 3015
QY 388 -----GlyHisArgAlaAlaValAsnVa 395
DB 3016 GAATACAGACTCAGTTCAAAGTGGCTGATTCGACAGAGACATTTAAGTTGGTTCATGG 3075
QY 395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLy 413
DB 3076 TGTGATGTTTCTCTGATGGATCATCTTTTGCATCTCTCTGATGACCAACATCAG 3135
QY 413 sValTyrAsnThr----- 417
DB 3136 GCTCTGGAGACAAAGAAAGATGTAAGAACTCTGCTGTAATGTTAAAGCAAGAGTAGA 3195
QY 417 ----- 417
DB 3196 TGTGTGTTTCAAGAAAATGAAGTATGCTGCTGCTGATGCTGATGACCATATAGACGCTGCA 3255
QY 418 -----SerThrCy 420
DB 3256 ACTCATTAATGGAAGAACAGGTCAAGATTGATTATCTGACTGAAGCTCAAGTACGCTGTG 3315
QY 420 s-----GluPh 422
DB 3316 TTGCTTAAGTCCACATCTTCAGTACATTCGATTTGGAGATGAAAATGAGCCATGAGAT 3375
QY 422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
DB 3376 TTTAGAACTGTAAACAATAGAACTCTTCAGCTCCAGGTTTCAGACACAAGAACTGATG 3435
QY 434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrI 452
DB 3436 GCATATCCAGTTCACAGCCGATGAGAGACTCTTTATTTCAAGTTCGTGATGCTGAAAT 3495
QY 452 eArgLeuTyrAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGlu 472
DB 3496 TCAGGTATGAATGGCAATTCGACAAATGTATC---TTTCTACAGAGCCATCAGGAAAC 3552
QY 472 uValArgCysIleArg-----PheAspAsnLysArgIleValSerGlyAlaTyrAspGly 491
DB 3553 AGTGAAAGACTTTAGACTCTTGAATAATTCAGAACTGCTTCTTCTTGGTTCATTTGATGAAC 3612
  
```

QY	2	AspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArg	21
Db	1769	GAAGAAGTTGAACACATCATGCAGGAG-----TTTGTAAATAAGCTCTCTTTTA	1816
QY	22	GluAspCysAsnAsnGlyGluProProArg-----LysIle	33
Db	1817	TTCTGTGATCGGATGGAAAGTCGTTTCGTTATTATTATCATGATCTTCAAGTAGATTTT	1876
QY	34	IleProGluLysAsnSerLeuArg-----	41
Db	1877	CTTACAGAGAGAATTCGACCCAGCTTCAGGATCTACATAAGACATAATCACTCACTTT	1936
QY	42	GlnThrTrpIrsenSerCysAlaArgLeuLysLeuAsnGlnGlu-ThrValCysLeuAlaIse	61
Db	1937	CAGAGATATCACGCCGCATCTCTTTACAGACTCAGGAAGCTGTATGTATGTGTAC	1996
QY	61	rThralaMethIysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe	81
Db	1997	AACITTT-CTGSCCTATCAC-----	2014
QY	81	rMetIleValProLysGlnArgLysLeuSerAlaSerTyrgluLysGluLysGluLeuCy	101
Db	2015	-----ATGCCCATGCCAAGATGCACAAGGAACCTTTG	2046
QY	101	eValLysTyrrPheGlu---GlnTrpSerGluSerAspGlnValGluPheVal-----	118
Db	2047	TGCTTTAATGTTTCCCTGCATTCGATTAAAGCA--AAACACAGAAGCTGTAGGCCCTGC	2103
QY	118	uHisIleIleSerGlnMetCysHisTyrglnHisGlyHisIleAsnSerTyrgluLysPr	138
Db	2104	TCATCTGATTCATGAATTTGTGGAAACACAGACAT-----AT	2139
QY	138	oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGl	158
Db	2140	ACTAGATGAAGAAGNTTGTCA-----GTCATGTA	2169
QY	158	uAsnIleLeuSerTyrgluAspAlaLys-----	167
Db	2170	GAATTTTCAGGAGTTTATTCTTTAAATGGACACACCTCTTGGACGACAGCCATTTCCTAA	2229
QY	168	-----SerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrr-----	180
Db	2230	TTATGTAACACTGGTCTCTGTGAGCGGAA-----ACTTCAGAAGTTTATCAGCAAGC	2283
QY	181	-----ArgValThrSerAspGlyMetLeuTrp-----Ly	190
Db	2284	TAAGCTGCAGGCCAAGCAGGAGGTGCGAATATGGAATGCTTTACCTGGTAATGATAAACA	2343
QY	190	slsLysLeuIleGluArgMetValarg-----ThrAspSerLeuTrpAr	204
Db	2344	AAAAAACATACAGAAATCTTTCCGCTTAGTGTGCGCCGCCACACAGATGCTGTIT-----	2398
QY	204	gGlyLeuAlaGluArgGlyTrpGlyGlnTyrgluPheLysAsnLysProProAspGl	224
Db	2398	-----	2398
QY	224	yAsnAlaProProAsnSerPheTyrgAlaLeuTyrrProLysIleIleGlnAspIleGl	244
Db	2399	-----TACCATGCCCTGCTTTCT-----GAGGATGTGCA	2427
QY	244	uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe	264
Db	2428	GAGATACCTTCT-----TGTGGAGCTGAATAAACCTTACAGGTGTC-----AAAGC	2475
QY	264	rGluThrSerLys-----GlyValTyrrCysLeuGl	274
Db	2476	TGAAACACAGGAGAGAACTTCTAGAAATCAAGGCTCATGAGATGAAGTGCCTTGTGTGC	2535
QY	274	nTyrr-----AspAspGlnLysIleValSerGlyLeuArgaspAsnThrIleLysIleTr	292
Db	2536	ATTCTCTACATGACAGATTTTATAGCACTGCTCAGTGGATAAAAAAGTAGAAGATTG	2595
QY	292	pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy	312

US-09-601-168B-2 (1-569) x US-09-435-115-1 (1-7042)

```

2596 GAATCTTATGACTGGGAACACTAGTACACACTATGATGAGCACTCAGACCAAGTCAATG 2655
312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerThr 328
2656 CTGCCATTTCACCAACAGTAGTCATCTCTCTTAGCCACTGGGCTCAAGTGACTGTT 2715
328 rValArgValThrAspValLeuThrGlyGluMetLeuAsnThrLeuIleHisCysG1 348
2716 CCTCAAACTTTGGGATTTGAATCAAAAAGAATGCGGAATACCATGTTTGGTCATACAA 2775
348 uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366
2776 TTCAGTCAATCACTGCAGATTTTCCACGATGATAAGCTTTTGGCTAGTTGTTGACGTGA 2835
366 pArgSerIleAlaValThrAspMetAlaSerProThrAsp-----IleThrLeuArg 383
2836 TGGAACTTAAGCTTTGGGATGCGCATCAGCAAAATGAGAGGAAAGCATTAATGTGAA 2895
383 g----- 383
2896 ACAGTCTCTCTAAATTTGGAGGACCTCAAGAGGATATGGAAGTATGATGAAAGTGTG 2955
384 -----ArgValLeuVal----- 387
2956 TTCGTGGTCTCTGATGTTGCGAAGGATAATGGTGCGAGCAAAATAAATAATCTTTTGTG 3015
388 -----GlyHisArgAlaAlaValAsnVa 395
3016 GAATACAGACTCAGCTTCAAAAGTGGGTGATGTCAGAGGACATTTAAGTTGGGTTCTGCG 3075
395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleu 413
3076 TGTGATGTTTCTCTGTGATGATCATCTTTTGACATCTCTGATGACGACGACATCAG 3135
413 sValThrAsnThr----- 417
3136 GCTCTGGGACACAAGAAAGTATGTAAAGACTCTGCTGTAATGTTAAAGCAAGATAGA 3195
417 ----- 417
3196 TGTGTGTTTCAAGAAAAATGAAGTATGATGTCCTTCAGTTGACCATATAGAGCTCTGCA 3255
418 -----SerThrCys 420
3256 ACTCATTAAATGAAGAACAGCTCAGATTGATTATCTGACTGAAGCTCAAGTTAGCTGTG 3315
420 g-----GluPh 422
3316 TTGCTTAAAGTCCACATCTTCAGTACATTGTCATTTGGAGATGAAAAATGAGCCCATGAGAT 3375
422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
3376 TTAGAACTTGTAAACAATAGAAATCTTCCAGTCCAGGTTTCAGCACAAGAAATCTGATG 3435
434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrI1 452
3436 GCACATCCAGTTTCACAGCCGATGAGAGACTCTTATTTCAGATTCTGATGATGCTGAAAT 3495
452 eArgLeuThrAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLe 472
3496 TCAGTATAGGAATGGCAATTTGGCAAAATGTATC-----TTTCTAGAGGCCATCAGAAAC 3552
472 uValArgCysIleArg-----PheAspAsnLysArgIleValSerGlyAlaTyrAspGly 491
3553 AGTGAAGAGACTTTAGACTCTTGAAAAATTCAGACTGCTTTCTTGGTCAATTTGATGAAAC 3612
491 sIleLysValThrAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCy 511
3613 AGTGAAGGATGGAAATATTATTCTGGAATAAAGAA-----LeuGlnPheAspG1 529
511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg----- 511
```

```

3650 -----AAAGACTTTGTCTGTACCAGGGTACAGTACTTCTTGTGACATTTCTCAGATGC 3705
529 uPheGlnIleValSerSerHisAspAspThrIleLeuIleTyrAspPhe 546
3706 TACCAAGTTTTCATCTCTGTGACAGAGACTGCAAGAGATCTGGAGTTT 3757

RESULT 11
US-09-098-310-1
; Sequence 1, Application US/09098310
; Patent No. 6403765
; GENERAL INFORMATION:
; APPLICANT: Alnemeti, Emad S.
; TITLE OF INVENTION: TRUNCATED APAF-1 AND METHODS OF USE
; FILE REFERENCE: 480140.438
; CURRENT APPLICATION NUMBER: US/09/098,310
; CURRENT FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (578)...(4159)
US-09-098-310-1

Alignment Scores:
Pred. No.: 1,59e-23 Length: 7042
Score: 299.50 Matches: 154
Percent Similarity: 33.72% Conservative: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 4 Gaps: 34

US-09-601-168B-2 (1-569) x US-09-098-310-1 (1-7042)
QY 2 AspProAlaGluAlaValLeuGlnGlyLysAlaLeuLysPheMetAsnSerSerGluArg 21
1769 GAAGAAGTTGAAGACATACATCGAGGAG-----TTGTAAATAGATCTCTTTTA 1816
22 GluAspCysAsnAsnGlyGluProArg-----LysIle 33
1817 TCTGTGATCGGAATGGAAGTCGTTTCTGTTATTATTACATGATCTTCAAGTAGATTT 1876
34 IleProGluLysAsnSerLeuArg----- 41
1877 CTTACAGAGAAAGTTGCGAGCCAGCTTCAGGATCTACATACAGATATACCTCAGTTT 1936
42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaSe 61
1937 CAGAGATATCACCAGCCGCTACTCTTTCCAGCATGACGAGAGACTGATGTTGGTAC 1996
61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
1997 AACTTT-CTGGCCTATCAC----- 2044
81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
2015 -----ATGCCAGATGCCAAGATGCAAGAACTTTG 2046
101 sValLysTyrPheGlu-----GlnTyrSerGluSerAspGlnValGluPheVal-----G1 118
2047 TGCTTTAATGTTTTCCTCGATGGATGGAATAAAGCA-----AAAACAGAACTTGTAGGCCCTGC 2103
118 uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
2104 TCATCTGATTCGAATTTGGAATACAGCAT-----AT 2139
138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaG1 158
2140 ACTAGATGAAAGGATTTGTGCA-----GTCAGTGA 2169
```

```

QY 158 uAsnIleuSerTyrIleuAspAlaLys----- 167
DB 2170 GAATTTTCAGGAGTTTATCTTTAAATGGACACCTTCCTGGAGGACAGCCATTCCTAA 2229
QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr----- 180
DB 2230 TATTGTACAACACTGGGTCTCTGTGAGCCGAA-----ACTTCAGAAGTTTATCAGCAAGC 2283
QY 181 -----ArgValThrSerAspGlyMetLeuTrp----- 190
DB 2284 TAAGCTGAGCCCAAGCAGGAGGTGCGATAATGGAAATGCTTTTACCTGGAATGGATAACAA 2343
QY 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTrpAr 204
DB 2344 AAAAAACATCAGGAATCTTTCCCGCTTAGTTGTGCGGCCCCACACAGATGCTGTT----- 2398
QY 204 gGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProAspGl 224
DB 2398 ----- 2398
QY 224 yAsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGl 244
DB 2399 -----TACCATGCTGCTTTTCT-----GAGATGCTCA 2427
QY 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
DB 2428 GAGAATAGCTTCT-----TGTGAGCTGTATAAAACCTTACAGGTGTC-----AAAGC 2475
QY 264 rGluThrSerLys-----GlyValTyrCysLeuGl 274
DB 2476 TGAACAGGAGAGAAACTTCTAGAAATCAAGGCTCATGAGATGAAGTGTGTTGTTGTGC 2535
QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
DB 2536 ATTCTCTACAGATGACAGATTATAGCAACCTGCTCAGTGTGATAAAAAAGTGAGATTG 2595
QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
DB 2596 GAATTTCTATGACTGGGAACCTAGTACACACCTATGATGAGCACTCAGAGCAAGTCAATTG 2655
QY 312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh 328
DB 2656 CTGCGCATTCACCAACAGTAGTCATCTCTCTTAGCCACTGGGTCAAGTCAAGTGCCTT 2715
QY 328 rValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGl 348
DB 2716 CCTCAAACTTTGGGATTGAATCAAAAAAGAAATGCGAAATACCATGTTTGGTCATACAAA 2775
QY 348 uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366
DB 2776 TTCAGTCAATCAGTCGAGATTTTCCAGCATGATGAAGCTTTTGGCTAGTTGTTCAGCTGA 2835
QY 366 pArgSerIleAlaValTrpAspMetAlaSerProThrAsp-----IleThrLeuAr 383
DB 2836 TGAACCTTAAAGCTTTGGGATGCGACATCAGCAATGAGAGGAAAGCATTAATGTGA 2895
QY 383 g----- 383
DB 2896 ACAGTTCTCTCTAAATTTGGAGGACCTCAAGAGGATATGGAAGTATAGTGAAGTGTG 2955
QY 384 -----ArgValLeuVal----- 387
DB 2956 TTCGTGGTCTGCTGATGGTGCAGGATTAAGTGGCAGCAAAAAATAAAATCTTTTGTG 3015
QY 388 -----GlyHisArgAlaAlaValAsnVa 395
DB 3016 GAATACAGACTCAGCTTCAAAAGGTGGCTGATGTCAGAGGACATTTAAGTTGGGTTCATGG 3075
QY 395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleY 413
DB 3076 TGTGATGTTTCTCTGATGATGATCATCATTTTTCACATCTTCTGATGACCAGACAATCAG 3135

```

```

QY 413 sValTrpAsnThr----- 417
DB 3136 CTTCTGGGAGCAAAAGAAATATGTAAGAACTCTGCTGTGTAATGTTAAAGCAAGAGTAGA 3195
QY 417 ----- 417
DB 3196 TGTGTGTTTTCAAGAAAATGAAGTGTGCTTTCAGCTTGGAGTATGACCATATAGACGCTCTGCA 3255
QY 418 -----SerThrCy 420
DB 3256 ACTCATTAATGGAAGAACAGGTCAGATTGATTAFTCTGACTGAAGCTCAAGTTAGCTGCTG 3315
QY 420 s-----GluPh 422
DB 3316 TTGCTTAAGTCCACATCTTCAGTACATTCATTTGGAGATGAAATGGAGCCATTGAGAT 3375
QY 422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
DB 3376 TTTAGAACTTGTAAACAATAGAAATCTTCCAGTCCAGGTTTCCAGCACAAAGAAACTGTATG 3435
QY 434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrI 452
DB 3436 GCATCCAGTTTCACGCCGATGGAAGACTCTTATTTCAAGTTCTGATGATGCTGAAAT 3495
QY 452 eArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluLe 472
DB 3496 TCAGTATGAATTTGGCAATTTGGCAAAATGTATC---TTTCTACGAGGCCATCAGGAAC 3552
QY 472 uValArgCysIleArg---PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLy 491
DB 3553 AGTGAAGACTTTAGACTCTTTGAAAAATTCAGAACTGCTTCTTGTGTCATTTGATGGAAC 3612
QY 491 sIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCy 511
DB 3613 AGTGAAGTATGGAATATTATTACTGGAATAAAGAA----- 3649
QY 511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspGl 529
DB 3650 ---AAAGACTTTCTCTGTCTCACCAGGTGACAGTACTTTCTTGTGACATTTCTCACCAGTGC 3705
QY 529 uPheGlnLeuValSerSerHisAspAspThrIleLeuIleTrpAspPhe 546
DB 3706 TACCAAGTTTTCATCTACCTCTGCTGCAAGACTGCAAAAGATCTGGAGTTT 3757

RESULT 12
US-09-690-364-21
; Sequence 21, Application US/09690364
; Patent No. 6468795
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
; FILE REFERENCE: RTS-0190
; CURRENT APPLICATION NUMBER: US/09/690,364
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 21
; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (578)...(4162)
US-09-690-364-21

Alignment Scores:
Pred. No.: 1,59e-23 Length: 7042
Score: 299.50 Matches: 154
Percent Similarity: 33.77% Conservative: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 4 Gaps: 34

```

US-09-601-168B-2 (1-569) x US-09-690-364-21 (1-7042)

```
QY 2 AspProAlaGluAlaValLeuGlnGluLysPheMetAsnSerSerGluArg 21
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1769 GAAGAAGTTGAAGACATACACTGACAGAG-----TTTGTAAATAAGTCTCTTTA 1816
QY 22 GluAspCysAsnAsnGlyGluProArg-----Lysile 33
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1817 TCTGTGATCGAATGGAAGTCTGTTGTTATTAATTATCATGATCTTCAAGTAGATTT 1876
QY 34 IleProGluLysAsnSerLeuArg----- 41
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1877 CTTACAGAGAAGAAATTGCAGCAGCAGCTTCAGGATCTACATAAGAAGATAATCACTCAGTTT 1936
QY 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaase 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1937 CAGACATATCACCAGCCCATCTCTTCCACCAGATCAGGAAGACTGTATGTTGGTAC 1996
QY 61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1997 AACITTT-CTGGCCTATCAC----- 2014
QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2015 -----ATGGCCAGTGGCCAGATGCACAGGAAGTTTG 2046
QY 101 sValLysTyrPheGlu---GlnTrpSerGluSerAspGlnValGluPheVal-----G1 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2047 TGCTTTAATGTTTCCCTCGGATTCGATTAAGCA---AAACAGAACTTGTAGGCCCTGC 2103
QY 118 uHisIleLeuSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2104 TCATCTGATTGATGATTTGGATACAGACAT-----AT 2139
QY 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaG1 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2140 ACTAGATGAAAAGGATTTGCA-----GTCAGTGA 2169
QY 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2170 GAATTTTCAGGAGTTTATCTTTAAATGGACACTTCTTGGACGACGCCATTTCCTAA 2229
QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTyrTrp----- 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2230 TATTGTACAAGTGGGTCTCTGAGCGCGAA-----ACTTCAGAAAGTTTATCAGCAAGC 2283
QY 181 -----ArgValThrSerAspGlyMetLeuTrp-----Lys 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2284 TAAGCTGCAGGCCAAGCAGGAGTGCATATGGAATGCTTTACTGGAATGGATAAACAA 2343
QY 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTrpAr 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2344 AAAAACAATCAGAACTCTTCCCGCTTAGTTGTCGCCGCCACACAGATGCTGTT----- 2398
QY 204 gGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrPheLeuPheLysAsnLysProAspG1 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2398 ----- 2398
QY 224 yAsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG1 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2399 -----TACCATGCTGCTTTCT-----GAGGATGCTCA 2427
QY 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGluArgIleHisCysArgSe 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2428 GAGATAGCTTCT-----TGTTGGAGCTGTATAAAACCTTACAGGTGTTTC---AARGC 2475
QY 264 rGluThrSerLys-----GlyValTyrCysLeuG1 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2476 TGAACAGGAGAGAACTTCAGAAATCAAGGCTCATGAGATGAAGTGTGTTGTTGTC 2535
QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2536 ATTCTCTACAGATGACAGATTATAGCAACCTGCTCATGTGATAAAAAAAGTGAAGATTG 2595
```

```

QY 511 sleuArgThrIleuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspG1 529
DB 3650 ----AAAGACTTTCTCTCCAGGGTACAGTACTTTCTTGACATTTCTCAGCATGC 3705
QY 529 uPheGlnIleValSerSerHisAspThrIleLeuIleTirPhe 546
DB 3706 TACCAAGTTTCACTCTCTGTGACAAAGCTGCAAGACTCTGGAGTTT 3757

RESULT 13
US-09-092-508-15
; Sequence 15, Application US/09092508
; Patent No. 6291643
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; STREET: 3100 No. 6291643west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092.508
; FILING DATE: 05-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,807
; FILING DATE: 05-JUN-1997
; APPLICATION NUMBER: 60/055,258.
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4192
; OTHER INFORMATION:
US-09-092-508-15

Alignment Scores:
(Pred. No.: 1.6e-23 Length: 7075
Score: 299.50 Matches: 154
Percent Similarity: 33.77% Conservative: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 3 Gaps: 34)

US-09-601-168B-2 (1-569) x US-09-092-508-15 (1-7075)

QY 2 AspProAlaGluAlaValLeuGlnGlnLysAlaLeuLysPheMetAsnSerSerGluArg 21
DB 1802 GAAGAAGTTGAAGACATCTCAGGAG-----TTGTAAATAAGTCTCTTTTA 1849
QY 22 GluAspCysAsnGlyGluProArg-----Lysile 33

```

```

DB 1850 TTCTGTGATCGGAATGGAAGTCTGTTGTTATTATTATACATGATCTTCAAGTAGATTTT 1909
QY 34 IleProGluLysAsnSerLeuArg----- 41
DB 1910 CTTACAGAGAAGAAATTCAGCCAGCTTCAGGATCTACATAAGAGATAATCACTCAGTTT 1969
QY 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaSe 61
DB 1970 CAGAGATATCCACCGCCGATCTTTTCACCATCAGGAGACTGTATGTATGTATG 2029
QY 61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
DB 2030 RACTTT-CTGGCCTATCAC----- 2047
QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
DB 2048 -----ATGCCAGATGCCAAGATGCCAAGAACTTTG 2079
QY 101 sValLysTyrPheGlu---GlnTirSerGluSerAspGlnValGluPheVal-----G1 118
DB 2080 TGCITTAATGTTTCCCTGGATTGGATTAAAGCA---AAACAGAACTTTAGSCCTGC 2136
QY 118 uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
DB 2137 TCATCTGATTCATGAATTTGTGGAATACAGACAT-----AT 2172
QY 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaG1 158
DB 2173 ACTAGTAAAGGATTGTGCA-----GTCAGTGA 2202
QY 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
DB 2203 GAATTTTCAGGAGTTTATCTTTAAATGGACACCTTCTTGACACAGCCATTTCTCTAA 2262
QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluIrrPtyr----- 180
DB 2263 TATTGTACAACCTGGGTCTCTGTAGCCGGAA-----ACTCAGAAAGTTTATCAGCAAGC 2316
QY 181 -----ArgValThrSerAspGlyMetLeuTirp-----Ly 190
DB 2317 TAAGCTGCAGGCCAAGCAGAGGTGCGATATGGAATGCTTTACCTGGAAATGGATAACAA 2376
QY 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTirPAr 204
DB 2377 AAAAAACATCACGAATCTTCCCGTTAGTTGTCGCCGCCACACAGATGCTGT----- 2431
QY 204 gGlyLeuAlaGluArgArgGlyTirPlyGlnTyrLeuPheLysAsnLysProAspG1 224
DB 2431 ----- 2431
QY 224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG1 244
DB 2432 -----TACCATGCTGCTTTCT-----GAGGATGTGCA 2460
QY 244 uThrIleGluSerAsnTirPArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
DB 2461 GAGAATAGCTTCT-----TGTGGAGCTGATAAAACCTTACAGGTGTTC---AAAGC 2508
QY 264 rGluThrSerLys-----GlyValTyrCysLeuG1 274
DB 2509 TGAACACAGAGAGAACTTCTAGAAATCAAGCTCATGAGGATGAAGTGTCTTTGTGTGC 2568
QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyr 292
DB 2569 ATTCTCTACAGATGACAGATTTATAGCAACCTGCTCAGTGTGATAAAAAGTGAAGATTG 2628
QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
DB 2629 GAATTCATGACTGGGAACTAGTACACACCTATGATGAGCAGCTCAGAGCAAGTCAATTG 2688
QY 312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh 328

```

```
Db      2689 CTGCCATTTTCAACCAAGTAGTCATCTCTCTTAGCCACTGGGTCAAGTGACTGCTT 2748
Qy      328 rValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysG1 348
Db      2749 CCTCAAACTTTGGGATTTGAATCAAAAAGATGTGCAAAATACCATGTTTGGTGCATACAAA 2808
Qy      348 uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366
Db      2809 TTCAGTCATCAATCGCAGATTTTACCAGATGATAGCTTTTGGCTAGTGTGTTCAGCTGA 2868
Qy      366 pArgSerIleAlaValTrpAspMetAlaSerProThrAsp-----IleThrLeuAr 383
Db      2869 TGGAACTTAAGCTTTGGGATCGCATCGCAAAATGAGAGAAAAGCAATTAATGTGAA 2928
Qy      383 g----- 383
Db      2929 ACAGTTCTTCTTAATTTGGAGACCTCTAGAGGATATGGAAGTATAGTGAAGTGTG 2988
Qy      384 -----ArgValLeuVal----- 387
Db      2989 TTCGTGCTGCTGATGGTCAAGGATAATGTGGCGAGCAAAAATAAAATCTTTTGTG 3048
Qy      388 -----GlyHisArgAlaAlaValAsnVa 395
Db      3049 GAATACAGATCAGTTCAAAGTGGCTGATTCAGAGGACATTTAAGTTGGTTCATCGG 3108
Qy      395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLy 413
Db      3109 TGTGATGTTTCTCTGATGATCATCATTTTGTGACATCTTGTGATGACCAACAATCAG 3168
Qy      413 sValTrpAsnThr----- 417
Db      3169 GCTCTGGAGACAAAGAAAGTATGTAAAGACTCTGCTGTAATGTTAAAGCAAGAAGTAGA 3228
Qy      417 ----- 417
Db      3229 TGTGTGTTTCAAGAAATCAAGTGATGGTCTTGCAGTTGACCATATAGAAGCGTCTCCA 3288
Qy      418 -----SerThrCy 420
Db      3289 ACTCATTAAATGGAAGAACAGGTGAGATTGATATCTGACTGAAGTCAAGTTAGCTGTG 3348
Qy      420 s-----GluPh 422
Db      3349 TTGCTTAAGTCCACATCTTCAGTACATTGCAATTGGAGATGAATAATGGAGCCATGAGAT 3408
Qy      422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
Db      3409 TTTAGAACTTGTAAACAATAGAATCTTCAGTCCAGTTTCAGCACACAGAAACTGTATG 3468
Qy      434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrIl 452
Db      3469 GCACATCCAGTTCACAGCCGATGAGAAGACTCTTATTCAAGTTCTGATGATGCTGAAAT 3528
Qy      452 eArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLe 472
Db      3529 TCAGGTATGAATTTGCAATTTGGCAAAATGTATC---TTTCTACGAGGGCCATCAGGAAC 3585
Qy      472 uValArgCysIleArg---PheAspAsnLysArgIleValSerGlyAlaValAspGlyLy 491
Db      3586 AGTGAAGACTTTAGACTCTTTGAAAATTCAGACTGCTTTCTTGGCTATTGATGGAAC 3645
Qy      491 sIleLysValTrpAspLeuValAlaLeuLeuAspProArgAlaProAlaGlyThrLeuCy 511
Db      3646 AGTGAAGTATGGAATATTATTACTGAAAATAAGAA----- 3682
Qy      511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspGl 529
Db      3693 ----AAAGACTTTCTGTCACCGGGTACAGTACTTTCTTGTGACATTTCTCAGATGC 3738
Qy      529 uPheGlnIleValSerSerHisAspThrIleLeuIleTrpAspPhe 546
Db      3739 TACCAAGTTTTCATCTACCTCTGCTGACAGACTGCAAAAGATCTGGAGTTT 3790
```

RESULT 14

```
US-09-435-115-15
; Sequence 15, Application US/09435115
; Patent No. 6346607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAP-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Merchant, Gould, Smith, Bdeil, Welter & Schmidt
; STREET: 3100 No. 6346607 West Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669, 60SU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4192
; OTHER INFORMATION:
; US-09-435-115-15
Alignment Scores:
Pred. No.: 1,6e-23 Length: 7075
Score: 299.50 Matches: 154
Percent Similarity: 33.77% Conservatism: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 4 Gaps: 34
US-09-601-168B-2 (1-569) x US-09-435-115-15 (1-7075)
Qy      2 AppProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArg 21
Db      1802 GAAGAAGTTGAAGACATCTGCGAG-----TTGTAAATAAGTCTCTTTTA 1849
Qy      22 GluAspCysAsnAspGlyGluProArg-----Lysile 33
Db      1850 TTCTGTGATCGGAATGGAAGTCGTTTCGTATTATTACATGATCTTCAAGTAGATTTT 1909
Qy      34 lleProGluLysAsnSerLeuArg----- 41
Db      1910 CTTACAGAGAGAAATTGCGAGCGCAGCTTCAGGATCTACATAGAAGATAATCACTCAGTTT 1969
```


QY 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaSe 61
 DB 1970 CAGAGATATCATCCAGCCGACATACCTTTCCACCAAGTACAGAAAGACTGTATGATGGTAC 2029
 QY 61 rThrAlaMetIysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
 DB 2030 AACITTT-CTGGCCTATCAC----- 2047
 QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
 DB 2048 -----ATGGCCAGTGCACCAAGATGCACCAAGCAACTTTG 2079
 QY 101 svallysTyrPheGlu---GlnTrpSerGluSerAspGlnValGluPheVal-----G1 118
 DB 2080 TGCTTTATGTTTCCCTGGATTGGATTGAAGCA---AAACAGAACTGTAGGCCCTGC 2136
 QY 118 uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
 DB 2137 TCATCTGATTCATGAATTTGTGGAATACAGACAT-----AT 2172
 QY 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGl 158
 DB 2173 ACTAGATGAAAGGATTGTGCA-----GTCAGTGA 2202
 QY 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
 DB 2203 GAATTTTCAGCAGTTTATCTTTAAATGGACACTTCTTGGACGACAGCCATTTCTCTAA 2262
 QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr----- 180
 DB 2263 TATTGTCAACTGGGTCTCTGTGACCGGAA-----ACTTCAGAACTTTATCAGCAAGC 2316
 QY 181 -----ArgValThrSerAspGlyMetLeuTrp-----Ly 190
 DB 2317 TAAGCTGCAGCCCAAGCAGAGGTCGATATGGAATGCTTTACCTGGAATGGATAACAA 2376
 QY 190 slsLeuIleGluArgMetValArg-----ThrAspSerLeuTrpPar 204
 DB 2377 AAAAACAATCACGAATCTTTCCCGCTTAGTTGTCGCCGCCCCACACAGATGCTGTT----- 2431
 QY 204 gGlyLeuAlaGluArgArgGlyTyrGlyTyrLeuPheLysAsnLysProProAspGl 224
 DB 2431 ----- 2431
 QY 224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGl 244
 DB 2432 -----TACCATGCCTGCTTTTCT-----GAGGATGGTCA 2460
 QY 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
 DB 2461 GAGAATAGCTTCT-----TGTTGAGCTGATAAAACCTTACAGTGTC---AAAGC 2508
 QY 264 rGluThrSerLys-----GlyValTyrCysLeuGl 274
 DB 2509 TGAACACAGAGAGAAACCTTCTAGAAATCAAGGCTCATGAGGATGAAGTGTGTTGTTGTC 2568
 QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
 DB 2569 ATTCCTACAGATGACAGATTATAGCAACCTGCTCAGTGGATAAAAAGTGAAGATTG 2628
 QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
 DB 2629 GAATTTCTATGACTGGGAACCTAGTACACACCTATGATGAGCACTCAGACGAAGTCAATTG 2688
 QY 312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh 328
 DB 2689 CTGCCATTTTCAACCAAGTAGTCATCTCTCTTGTGCCACTGGGTCAAGTGAAGTCTT 2748
 QY 328 rValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGl 348
 DB 2749 CTCAAACTTTGGGATTTGAATCAAAAAGATGTGAAATACCATGTGTTGGTCATACAAA 2808
 QY 348 uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366

DB 2809 TTCACTCAATCACTGCGATTTTCCACAGATGATAAGCTTTTGGCTAGTGTGTTCACTGA 2868
 QY 366 pArgSerIleAlaValTrpAspMetAlaSerProThrAsp-----IleThrLeuArg 383
 DB 2869 TGGAACTTTAAAGCTTTGGGATCGGACATCAGCAAAATGAGAGGAAAAAGCATTAAATGTGAA 2928
 QY 383 g----- 383
 DB 2929 ACAGTTCTTCTAAATTTGGAGACCCCTCAAGAGGATATGGAAGTGATAGTGAAGTTTG 2988
 QY 384 -----ArgValLeuVal----- 387
 DB 2989 TTCGTGCTCTGCTGATGGTCAAGGATAATGTGGCGACAAAAATAAATCTTTTGTG 3048
 QY 388 -----GlyHisArgAlaAlaValAsnVa 395
 DB 3049 GAATACAGACTCAGTTCAAAGGTGGCTGATTCAGAGGACATTTTAAGTTGGTTTCATGG 3108
 QY 395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLy 413
 DB 3109 TGTGATGTTTCTCTGATGGATCATCTTTTGACATCTTCTGATGACACACATCAG 3168
 QY 413 sValTrpAsnThr----- 417
 DB 3169 GCTCTGGGAGACAAAGAAAGATATGTAAGAACTCTGCTGTAATGTTAAAGCAAGAGTAGA 3228
 QY 417 ----- 417
 DB 3229 TGTGTGTTTCAAGAAAAATGAAGTATGTCCTTCGAGTTGACCATATAAGACGTCTGCA 3288
 QY 418 -----SerThrCy 420
 DB 3289 ACTCATTAATGGAAGACAGTCAGATGATGATATCTCACTGAAAGCTCAAGTTAGCTGCTG 3348
 QY 420 s-----GluPh 422
 DB 3349 TTGCTTAAGTCCACATCTTCAGTACATTCATGATTTGGAGATGAAATGGAGCCATTGAGAT 3408
 QY 422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
 DB 3409 TTTAGAACTTTGTAACAATAGAATCTTCCAGTCCAGGTTTCAGCACAAAGAAACTGTATG 3468
 QY 434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrIle 452
 DB 3469 GCATCCAGTTCACACCGATGAGAAGACTTATTTCAAGTTCTGATGATGCTGCTGAAAT 3528
 QY 452 eArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLe 472
 DB 3529 TCAGGTATGGAATGGCAATGGCAAAATGATC---TTTCTACGAGGCCATCAGGAAAC 3585
 QY 472 uValArgCysIleArg---PheAspAsnLysArgIleValSerGlyAlaLysArgGlyLy 491
 DB 3586 AGTGAAGACTTTAGACTCTTGAAAAATTCAGACTGCTTTCTTGTGCTATTTGATGGAAC 3645
 QY 491 sIleLysValTrpAspLeuValAlaLeuAspProArgAlaProAlaGlyThrLeuCy 511
 DB 3646 AGTGAAGTATGGAATATTTACTGGAATAAGAA----- 3682
 QY 511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspGl 529
 DB 3683 ---AAAGACTTTGCTGTCCAGGAGTACAGTACTTTCTTTGTGACATTTCTCAGCATGC 3738
 QY 529 uPheGlnIleValSerSerHisAspAspThrIleLeuIleTrpAspPhe 546
 DB 3739 TACCAAGTTTTCATCTCTGCTGACAGACTGCAAGATCTGGAGTTTT 3790

RESULT 15

US-08-188-582-17
 ; Sequence 17, Application US/08188582
 ; Patent No. 5534410
 ; GENERAL INFORMATION:
 ; APPLICANT: Tjian, Robert

Search completed: October 22, 2003, 17:13:31
Job time : 149 secs